Appendix A
Abbreviations and Notations

For a better overview, we collect here some abbreviations and specific notation.

Abbreviations

ala	Algebraic approach
ana	Analytical approach
ClM	Classical mechanics
CONS	Complete orthonormal system
CSCO	Complete system of commuting observables
DEq	Differential equation
EPR	Einstein-Podolsky-Rosen paradox
fapp	'Fine for all practical purposes'
MZI	Mach–Zehnder interferometer
ONS	Orthonormal system
PBS	Polarizing beam splitter
QC	Quantum computer
QM	Quantum mechanics
QZE	Quantum Zeno effect
SEq	Schrödinger equation

Operators

There are several different notations for an operator which is associated with a physical quantity A; among others: (1) A, that is the symbol itself, (2) $\hat{A}$, notation with hat (3) $\mathcal{A}$, calligraphic typeface, (4) A_{op}, notation with index. It must be clear from the context what is meant in each case.

For special quantities such as the position x, one also finds the uppercase notation X for the corresponding operator.

Many-Particle States

For two quantum objects, the position gives the object number, if nothing is otherwise specified:

$$|nm\rangle = |n_1 m_2\rangle ; \tag{A.1}$$

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

n and m each stand for a single or for several quantum numbers.

With more than two quantum objects (object 1 with quantum numbers α_1, object 2 with quantum numbers α_2)..., we generally use the following notation:

$$|1 : \alpha_1, 2 : \alpha_2, \ldots, n : \alpha_n\rangle \,. \tag{A.2}$$

It is more transparent than the equivalent notation

$$\left|\varphi^{(1)}_{\alpha_1}\varphi^{(2)}_{\alpha_2}\ldots\varphi^{(n)}_{\alpha_n}\right\rangle . \tag{A.3}$$

Interchanging the quantum numbers (e.g. those of object 1 and 2) looks like this:

$$|1 : \alpha_2, 2 : \alpha_1, \ldots, n : \alpha_n\rangle \tag{A.4}$$

instead of

$$\left|\varphi^{(1)}_{\alpha_2}\varphi^{(2)}_{\alpha_1}\ldots\varphi^{(n)}_{\alpha_n}\right\rangle . \tag{A.5}$$

The Hamiltonian and the Hadamard transformation

We denote the Hamiltonian by H. With reference to questions of quantum information, especially in Chap. 27, Vol. 2, H stands for the Hadamard transformation.

Perturbation Calculations

To denote Hamiltonians and states in perturbation theory, we use a superscript index in parentheses, indicating the perturbation order:

$$H^{(0)}; \;\; \left|\varphi^{(1)}\right\rangle \text{ etc.} \tag{A.6}$$

Tracing Out

The reduced density operator, arising through tracing out all degrees of freedom $\neq k$, is denoted by a superscript index in parentheses:

$$\rho^{(k)} \tag{A.7}$$

Vector Spaces

We denote a vector space by $\mathcal{V}$, a Hilbert space by $\mathcal{H}$.

Based on the notation $\mathbb{R}^3$ or $\mathbb{C}^3$ for the three-dimensional real or complex space, we select the following notation, if necessary, for a more precise specification of Hilbert spaces:

$$\mathcal{H}^d_{n(m)} \text{ with } \begin{aligned} d &= \text{dimension} \\ n &= \text{number of the corresponding quantum object} \\ m &= \text{total number of quantum objects.} \end{aligned} \tag{A.8}$$

Appendix B
Special Functions

We compile here some material for important special functions of quantum mechanics.

B.1 Spherical Harmonics

The general form of the *spherical harmonics* $Y_l^m(\vartheta, \varphi)$[1] is

$$Y_l^m(\vartheta, \varphi) = f_l^m(\vartheta)\, e^{im\varphi} = (-1)^{\frac{m+|m|}{2}} \left[\frac{2l+1}{4\pi} \frac{(l-|m|)!}{(l+|m|)!}\right]^{1/2} P_l^{|m|}(\cos\vartheta)\, e^{im\varphi}, \tag{B.1}$$

where P_l^m are the *associated Legendre functions*:

$$P_l^m(x) = \frac{(-1)^m}{2^l l!} \left(1-x^2\right)^{\frac{m}{2}} \frac{\mathrm{d}^{l+m}}{dx^{l+m}} \left(x^2-1\right)^l. \tag{B.2}$$

These are solutions of the differential equation

$$\left(1-x^2\right) \frac{\mathrm{d}^2 g(x)}{\mathrm{d}x^2} - 2x \frac{\mathrm{d}g(x)}{\mathrm{d}x} + \left[l(l+1) - \frac{m^2}{1-x^2}\right] g(x) = 0. \tag{B.3}$$

In particular, we have for $m = 0$ the *Legendre polynomials* $P_l(\cos\vartheta)$:

$$Y_l^0(\vartheta, \varphi) = \sqrt{\frac{2l+1}{4\pi}} P_l(\cos\vartheta). \tag{B.4}$$

The spherical harmonics form a CONS, they are complete:

[1] Also called spherical functions, surface spherical harmonics, Laplace spherical harmonics, Laplace spherical functions or the like.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_l^m(\vartheta,\varphi)\,Y_l^{m*}\left(\vartheta',\varphi'\right) = \frac{\delta\left(\vartheta-\vartheta'\right)\delta\left(\varphi-\varphi'\right)}{\sin\vartheta} \tag{B.5}$$

and orthonormal

$$\int_0^{\pi} \mathrm{d}\vartheta \sin\vartheta \int_0^{2\pi} \mathrm{d}\varphi Y_l^{m*}(\vartheta,\varphi)\,Y_{l'}^{m'}(\vartheta,\varphi) = \delta_{ll'}\delta_{mm'}. \tag{B.6}$$

With the notation $\Omega = (\vartheta, \varphi)$ for the solid angle and $\mathrm{d}\Omega = \sin\vartheta\, \mathrm{d}\vartheta\, \mathrm{d}\varphi$ (also written $\mathrm{d}^2\hat{r}$ or $\mathrm{d}\hat{r}$), the orthogonality relation is written as:

$$\int Y_l^{m*}(\vartheta,\varphi)\,Y_{l'}^{m'}(\vartheta,\varphi)\,\mathrm{d}\Omega = \delta_{ll'}\delta_{mm'}. \tag{B.7}$$

Thus, for the Legendre polynomials:

$$\int \mathrm{d}\Omega\, P_l(\cos\vartheta)\;P_{l'}(\cos\vartheta) = \frac{4\pi}{2l+1}\delta_{ll'}. \tag{B.8}$$

The addition theorem of the spherical harmonics reads

$$\frac{2l+1}{4\pi}P_l(\cos\alpha) = \sum_{m=-l}^{l} Y_l^{m*}(\vartheta_1,\varphi_1)\,Y_l^m(\vartheta_2,\varphi_2), \tag{B.9}$$

where α is the angle between the directions (ϑ_1, φ_1) and (ϑ_2, φ_2).

The product of two spherical harmonics is given by[2]:

$$\begin{aligned} Y_{l_1}^{m_1}(\boldsymbol{\vartheta},\varphi)\,Y_{l_2}^{m_2}(\boldsymbol{\vartheta},\varphi) = &\sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi}} \\ &\times \sum_{L=|l_1-l_2|}^{l_1+l_2}\sqrt{\frac{1}{2L+1}}\,\langle l_1 l_2 00\,|L0\rangle \\ &\times \langle l_1 l_2 m_1 m_2\,|L\,M\rangle\,Y_L^M\left(\hat{\mathbf{r}}\right);\; M = m_1+m_2. \end{aligned} \tag{B.10}$$

Finally, we give explicitly the first few spherical harmonics:

---

[2]For the proof, one uses properties of the rotation matrices which are related to the spherical harmonics by $D_{m0}^{(l)}(\vartheta,\varphi) = \sqrt{\frac{4\pi}{2l+1}}Y_l^{m*}(\vartheta,\varphi)$. For the Clebsch–Gordan coefficients see Chap. 16.

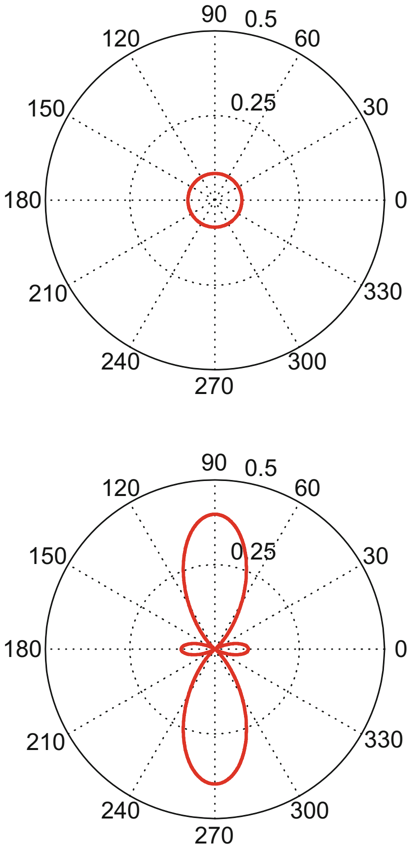


Fig. B.1 Polar diagram of $\left|Y_l^0\right|^2$ for $l = 0, 1, 2, 3$

$$Y_0^0 = \frac{1}{\sqrt{4\pi}};\ Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\vartheta;\ Y_1^1 = -\sqrt{\frac{3}{8\pi}}\sin\vartheta e^{i\varphi}$$
$$Y_2^0 = \sqrt{\frac{5}{16\pi}}\left(3\cos^2\vartheta - 1\right);\ Y_2^1 = -\sqrt{\frac{15}{8\pi}}\sin\vartheta\cos\vartheta e^{i\varphi};\ Y_2^2 = \sqrt{\frac{15}{32\pi}}\sin^2\vartheta e^{2i\varphi}$$
$$Y_3^0 = \sqrt{\frac{7}{16\pi}}\left(5\cos^3\vartheta - 3\cos\vartheta\right);\ Y_3^1 = -\sqrt{\frac{21}{64\pi}}\sin\vartheta\left(5\cos^2\vartheta - 1\right)e^{i\varphi}$$
$$Y_3^2 = \sqrt{\frac{105}{32\pi}}\sin^2\vartheta\cos\vartheta e^{2i\varphi};\ Y_3^3 = -\sqrt{\frac{35}{64\pi}}\sin^3\vartheta e^{3i\varphi}. \tag{B.11}$$

The graphical representation of some spherical harmonics is shown in Fig. B.1.

B.2 Spherical Bessel Functions

The stationary SEq for free quantum objects in spherical coordinates (cf. Chap. 17) reads:

$$E\psi = -\frac{\hbar^2}{2m}\nabla^2\psi = -\frac{\hbar^2}{2m}\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\mathbf{l}^2}{\hbar^2 r^2}\right)\psi. \tag{B.12}$$

Correspondingly, we can use the *ansatz* $\psi(\mathbf{r}) = y_l(r)\,Y_l^m(\vartheta,\varphi)$ and obtain

$$Ey_l(r) = -\frac{\hbar^2}{2m}\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{l(l+1)}{r^2}\right)y_l(r), \tag{B.13}$$

or

$$\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r + k^2 - \frac{l(l+1)}{r^2}\right)y_l(r) = 0;\; k^2 = \frac{2m}{\hbar^2}E. \tag{B.14}$$

The *spherical Bessel functions* are special solutions of these equations. The solutions which are regular at $r = 0$ are called proper spherical Bessel functions j_l; the irregular solutions are the Neumann functions n_l. Combinations of these functions are the Hankel functions $h_l^{(\pm)}$ of the first $(+)$ and second $(-)$ kind:

$$h_l^{(\pm)} = n_l \pm i j_l. \tag{B.15}$$

The functions with $l = 0$ and $l = 1$ are:

$$\begin{aligned} j_0 &= \frac{\sin kr}{kr};\; j_1 = \frac{\sin kr}{(kr)^2} - \frac{\cos kr}{kr} \\ n_0 &= \frac{\cos kr}{kr};\; n_1 = \frac{\cos kr}{(kr)^2} + \frac{\sin kr}{kr}. \end{aligned} \tag{B.16}$$

Functions with higher indices can be computed recursively; with $x = kr$, it follows for instance:

$$(2l+1)\,f_l = x\,(f_{l+1} + f_{l-1})\,;\; l \neq 0 \tag{B.17}$$

or

$$f_l = -x^{l-1}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{f_{l-1}}{x^{l-1}}\right) = \left[x^l\left(-\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^l\right]f_0. \tag{B.18}$$

Here, $f_l = c_1 j_l + c_2 n_l$ is an arbitrary linear combination of j_l and n_l.

Their behavior at the origin is given by

$$\left.\begin{aligned} j_l(x) &\sim \frac{x^l}{(2l+1)!!}\left[1+O\left(x^2\right)\right] \\ n_l(x) &\sim \frac{(2l+1)!!}{(2l+1)}\left(\frac{1}{x}\right)^{l+1}\left[1+O\left(x^2\right)\right] \end{aligned}\right\} \quad ; \quad x \to 0, \tag{B.19}$$

and the asymptotic forms are

$$\left.\begin{aligned} j_l(x) &\sim \frac{1}{x}\sin\left(x-\frac{l\pi}{2}\right)\left[1+O\left(\frac{1}{x}\right)\right] \\ n_l(x) &\sim \frac{1}{x}\cos\left(x-\frac{l\pi}{2}\right)\left[1+O\left(\frac{1}{x}\right)\right] \end{aligned}\right\} \; ; \; x \to \infty. \tag{B.20}$$

The spherical Bessel functions play an important role in (among others) scattering theory, since under certain conditions they constitute the asymptotic solutions. We have e.g. for the outgoing scattered wave:

$$\psi_{\text{out}} \to h_l^{(+)}(x) = n_l(x) + i j_l(x) \to \frac{e^{ix}}{x}. \tag{B.21}$$

Due to the relation

$$\begin{aligned} j_l(x) &= \sqrt{\frac{\pi}{2x}} J_{l+\frac{1}{2}}(x) \\ n_l(x) &= (-1)^l \sqrt{\frac{\pi}{2x}} J_{-l-\frac{1}{2}}(x), \end{aligned} \tag{B.22}$$

where $J_\nu(x)$ are the ordinary Bessel functions of order ν, the spherical Bessel functions are also called 'half-integer Bessel functions' or 'small Bessel functions'.

B.3 Eigenfunctions of the Hydrogen Atom

The potential is given by

$$V(r) = -\frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r} \tag{B.23}$$

where Ze is the nuclear charge and e the electronic charge. The eigenfunctions are

$$\psi_{nlm}(r) = R_{nl}(r)\, Y_l^m(\vartheta, \varphi). \tag{B.24}$$

Here, $n = 1, 2, \ldots$ is the principal quantum number, l and m determine the angular momentum.

The *radial functions* are given by:

$$R_{nl} = \sqrt{\frac{(n-l-1)!\,(2\kappa)^3}{2n\,((n+l)!)^3}}\,(2\kappa r)^l\, e^{-\kappa r} L_{n+l}^{2l+1}(2\kappa r) \tag{B.25}$$

with

$$\kappa = \frac{Z}{na_0}. \tag{B.26}$$

The radius a_0 is given by (μ = reduced mass):

$$a = \frac{a_0}{Z} = \frac{\hbar^2}{Z\mu e^2};\ a_0 \approx \text{ Bohr radius } a_B = \frac{\hbar^2}{me^2}. \tag{B.27}$$

The functions $L_{n+l}^{2l+1}(y)$ are the *associated Laguerre polynomials*; they can be calculated from

$$L_r^s(y) = \left(-\frac{\mathrm{d}}{\mathrm{d}y}\right)^s e^y \left(\frac{\mathrm{d}}{\mathrm{d}y}\right)^r e^{-y} y^r. \tag{B.28}$$

The first radial functions are:

$$\begin{aligned}
&\text{K-shell, s-orbital: } R_{10}(r) = 2\left(\tfrac{Z}{a_0}\right)^{\frac{3}{2}} e^{-\frac{Zr}{a_0}}\\
&\text{L-shell, s-orbital: } R_{20}(r) = 2\cdot\left(\tfrac{Z}{2a_0}\right)^{\frac{3}{2}}\left(1-\tfrac{Zr}{2a_0}\right) e^{-\frac{Zr}{2a_0}}\\
&\text{L-shell, p-orbital: } R_{21}(r) = \tfrac{1}{\sqrt{3}}\cdot\left(\tfrac{Z}{2a_0}\right)^{\frac{3}{2}} \tfrac{Zr}{a_0} e^{-\frac{Zr}{2a_0}}\\
&\text{M-shell, s-orbital: } R_{30}(r) = 2\cdot\left(\tfrac{Z}{2a_0}\right)^{\frac{3}{2}}\left(1-\tfrac{2Zr}{3a_0}+\tfrac{2Z^2r^2}{27a_0^2}\right) e^{-\frac{Zr}{3a_0}}\\
&\text{M-shell, p-orbital: } R_{31}(r) = \tfrac{4\sqrt{2}}{3}\cdot\left(\tfrac{Z}{3a_0}\right)^{\frac{3}{2}} \tfrac{Zr}{a_0}\left(1-\tfrac{Zr}{6a_0}\right) e^{-\frac{Zr}{3a_0}}\\
&\text{M-shell, d-orbital: } R_{32}(r) = \tfrac{2\sqrt{2}}{27\sqrt{5}}\cdot\left(\tfrac{Z}{3a_0}\right)^{\frac{3}{2}}\left(\tfrac{Zr}{a_0}\right)^2 e^{-\frac{Zr}{3a_0}}.
\end{aligned} \tag{B.29}$$

A graphical representation of some radial functions is shown in Figs. B.2 and B.3.

For the mean values in the state ψ_{nlm}, we have

$$\begin{aligned}
&\left\langle\frac{1}{r}\right\rangle = \frac{Z}{a_0 n^2};\ \left\langle\frac{1}{r^2}\right\rangle = \frac{Z^2}{a_0^2 n^3}\frac{1}{l+\frac{1}{2}};\ \left\langle\frac{1}{r^3}\right\rangle = \frac{Z^3}{a_0^3 n^3}\frac{1}{l\left(l+\frac{1}{2}\right)(l+1)}\\
&\langle r\rangle = \frac{1}{2}\left[3n^2 - l(l+1)\right]\frac{a_0}{Z};\ \left\langle r^2\right\rangle = \frac{1}{2}\left[5n^2+1-3l(l+1)\right] n^2\frac{a_0^2}{Z^2},
\end{aligned} \tag{B.30}$$

and for $s > -2l-1$, we find the recursion relation

$$\frac{s+1}{n^2}\left\langle r^s\right\rangle - (2s+1)\frac{a_0}{Z}\left\langle r^{s-1}\right\rangle + \frac{s}{4}\left[(2l+1)^2 - s^2\right]\frac{a_0^2}{Z^2}\left\langle r^{s-2}\right\rangle = 0. \tag{B.31}$$

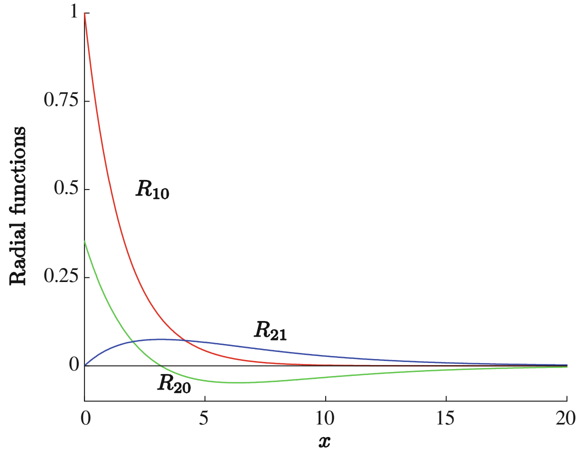


Fig. B.2 The radial functions R_{10} (*red*), R_{20} (*green*) and R_{21} (*blue*). Not normalized; $x = \frac{Zr}{a_0}$

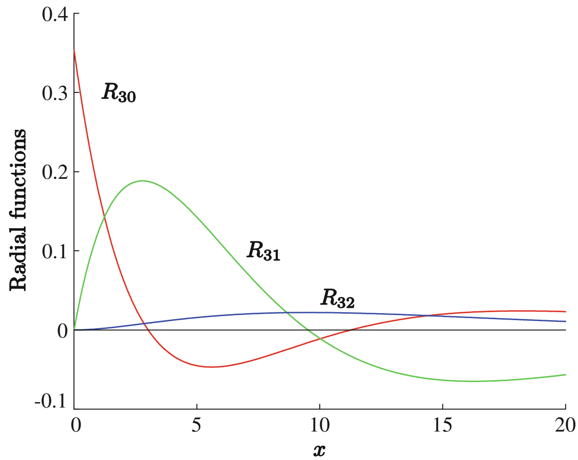


Fig. B.3 The radial functions R_{30} (*red*), R_{31} (*green*) and R_{32} (*blue*). Not normalized; $x = \frac{Zr}{a_0}$

B.4 Hermite Polynomials

The eigenfunctions of the harmonic oscillator can be written as

$$\psi_n(x) = \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} \frac{1}{\sqrt{n! \cdot 2^n}} \cdot e^{-\frac{x^2}{2}} \cdot H_n(x), \tag{B.32}$$

where the *Hermite polynomials* are defined by

$$H_n(x) = e^{\frac{x^2}{2}} \left(x - \frac{\mathrm{d}}{\mathrm{d}x}\right)^n e^{-\frac{x^2}{2}}. \tag{B.33}$$

The first few Hermite polynomials are

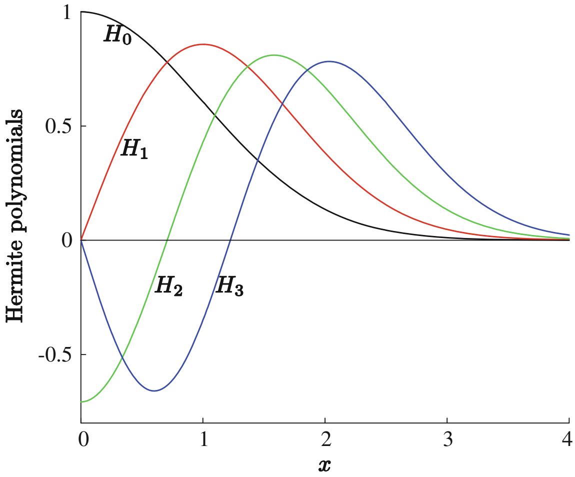


Fig. B.4 The functions $e^{-\frac{x^2}{2}} H_n(x)/\sqrt{n!2^n}$ for $n = 0, 1, 2, 3$ (*black, red, green, blue*)

$$H_0 = 1;\; H_1 = 2x;\; H_2 = 4x^2 - 2;\; H_3 = 8x^3 - 12x. \tag{B.34}$$

Further polynomials are best calculated recursively from

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x). \tag{B.35}$$

The Hermite polynomials, which belong to the important class of *orthogonal polynomials*, obey the orthogonality relation:

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} H_n(x) H_m(x) dx = \sqrt{\pi} n! 2^n \delta_{nm}. \tag{B.36}$$

A graphical representation of some Hermite polynomials is found in Fig. B.4.

B.5 Waves

A *plane wave*[3] travelling in the z direction (i.e. $\mathbf{k} = (0, 0, k)$) can be decomposed into partial waves:

$$e^{ikz} = \sum_{l=0}^{\infty} (2l + 1)\, i^l\, j_l\,(kr)\, P_l\,(\cos\vartheta)\,. \tag{B.37}$$

[3] Wave, because one considers explicitly only the spatial part of $e^{i(kz-\omega t)}$, adding tacitly the term $e^{-i\omega t}$.

Here, spherical coordinates are assumed.

Generalizing, we have for $\mathbf{r} \to (r, \vartheta_r, \varphi_r)$ and $\mathbf{k} \to (k, \vartheta_k, \varphi_k)$ the representation

$$e^{i\mathbf{k}\cdot\mathbf{r}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} i^l j_l(kr) Y_l^{m*}(\vartheta_k, \varphi_k) Y_l^m(\vartheta_r, \varphi_r) . \tag{B.38}$$

With the help of the addition theorem (B.9), we can also write:

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{l=0}^{\infty} (2l + 1) i^l j_l(kr) P_l(\cos\alpha) , \tag{B.39}$$

where α is the angle between the directions (ϑ_k, φ_k) and (ϑ_r, φ_r).

For *outgoing spherical waves*, we have (for incoming waves correspondingly $e^{-i\cdots}$):

$$\frac{e^{ik|\mathbf{r}_1-\mathbf{r}_2|}}{|\mathbf{r}_1 - \mathbf{r}_2|} = k \sum_{l=0}^{\infty} (2l + 1) i^l j_l(kr_<) h_l^{(+)}(kr_>) P_l(\cos\alpha) , \tag{B.40}$$

where α is the angle between the directions of $\mathbf{r}_1$ and $\mathbf{r}_2$, and the abbreviations are defined as $r_< = \min(r_1, r_2)$ and $r_> = \max(r_1, r_2)$. In particular, for $k = 0$ we have

$$\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} P_l(\cos\alpha) = \sum_{l,m} \frac{r_<^l}{r_>^{l+1}} \frac{4\pi}{2l + 1} Y_l^m(\vartheta_1, \varphi_1) Y_l^{m*}(\vartheta_2, \varphi_2) . \tag{B.41}$$

Appendix C
Tensor Product

We discuss here some properties of the *tensor product* of vector spaces. In order to make it more familiar, we will write down some results in explicit form (and in doing so, we will see that this notation is quite clumsy).

C.1 Direct Product

The tensor product connects two or more vector spaces to form a common vector space, also called the *product space*. Assuming two vector spaces $\mathcal{V}_1$ (dimension N) and $\mathcal{V}_2$ (dimension M), we write for the vector product $\mathcal{V}$

$$\mathcal{V} = \mathcal{V}_1 \otimes \mathcal{V}_2 \tag{C.1}$$

If the spaces $\mathcal{V}_1$ and $\mathcal{V}_2$ have the bases $\{|n\rangle_1\}$ and $\{|m\rangle_2\}$, the basis of the product space is the set of all pairs $\{|n\rangle_1 \otimes |m\rangle_2\}$. Thus, the dimension of the product space is $N \cdot M$, and a general state vector has the form

$$|\psi\rangle = \sum_{n,m} c_{nm} |n\rangle_1 \otimes |m\rangle_2 , \tag{C.2}$$

where we write also $|n_1 \otimes m_2\rangle$ or simply $|n\rangle_1 |m\rangle_2$, $|n_1\rangle |m_2\rangle$ or $|n_1 m_2\rangle$ instead of $|n\rangle_1 \otimes |m\rangle_2$. If the meaning is clear from context, we will omit the indices.

Here, we give an example in component representation: Given that

$$|u\rangle \cong \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}; \quad |v\rangle \cong \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}; \tag{C.3}$$

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

then we have[4]:

$$|w\rangle = |u\rangle \otimes |v\rangle \cong \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \otimes \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u_1 v_1 \\ u_1 v_2 \\ u_2 v_1 \\ u_2 v_2 \\ u_3 v_1 \\ u_3 v_2 \end{pmatrix}. \tag{C.4}$$

C.2 Direct Sum of Vector Spaces

We mention this term briefly since it is sometimes confused with the direct product. For the vector spaces $\mathcal{V}_1$ (dimension N) and $\mathcal{V}_2$ (dimension M), the direct sum is written as:

$$\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2, \tag{C.5}$$

where the space has dimension $N + M$. If $\mathcal{V}_1$ and $\mathcal{V}_2$ have the bases $\{|n\rangle_1\}$ and $\{|m\rangle_2\}$, then the basis of the sum space is the set $\{|1\rangle_1, |2\rangle_2, \ldots |1\rangle_2, |2\rangle_2, \ldots\}$.

Example in component representation: we have

$$|w\rangle = |u\rangle \oplus |v\rangle \cong \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \oplus \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ v_1 \\ v_2 \end{pmatrix}. \tag{C.6}$$

C.3 Properties of the Tensor Product

Tensor products can be carried out multiply; an example is:

$$\mathcal{V} = \mathcal{V}_1 \otimes \mathcal{V}_2 \otimes \cdots \otimes \mathcal{V}_n = \bigotimes_{l=1}^{n} \mathcal{V}_l. \tag{C.7}$$

For identical spaces ($\mathcal{V}_1 = \mathcal{V}_2 = \cdots$), one can write this as *tensor power* $\mathcal{V} = \mathcal{V}_1^{\otimes N}$ or shortly $\mathcal{V}_1^N$.

We assume for the following that the operator U acts only in space 1 and V only in space 2.

A tensor product of operators acts on a tensor product of vectors in each space separately:

[4]Rule: The right index changes the fastest.

$$(U \otimes V)\,(|u\rangle \oplus |v\rangle) \;=\; U\,|u\rangle \oplus V\,|v\rangle\,. \tag{C.8}$$

In contrast to proper operator products, the order is not changed in the adjoint

$$(U \otimes V)^{\dagger} \;=\; U^{\dagger} \otimes V^{\dagger};\;\; (|u\rangle \oplus |v\rangle)^{\dagger} \;=\; \langle u| \otimes \langle v|\,. \tag{C.9}$$

We have for example

$$\{(U_1U_2U_3) \otimes (V_1V_2)\}^{\dagger} \;=\; (U_1U_2U_3)^{\dagger} \otimes (V_1V_2)^{\dagger} \;=\; \left(U_3^{\dagger}U_2^{\dagger}U_1^{\dagger}\right) \otimes \left(V_2^{\dagger}V_1^{\dagger}\right). \tag{C.10}$$

C.4 Examples

C.4.1 General Examples

Given two matrices A and B with

$$A \;=\; \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix};\; B \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{C.11}$$

Then it follows for $A \otimes B$:

$$A \otimes B \;=\; \begin{pmatrix} 1B & 2B & 3B \\ 4B & 5B & 6B \\ 7B & 8B & 9B \end{pmatrix} \;=\; \begin{pmatrix} a & b & 2a & 2b & 3a & 3b \\ c & d & 2c & 2d & 3c & 3d \\ 4a & 4b & 5a & 5b & 6a & 6b \\ 4c & 4d & 5c & 5d & 6c & 6d \\ 7a & 7b & 8a & 8b & 9a & 9b \\ 7c & 7d & 8c & 8d & 9c & 9d \end{pmatrix} \tag{C.12}$$

and for $A \oplus B$

$$A \oplus B \;=\; \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \;=\; \begin{pmatrix} 1 & 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 & 0 \\ 7 & 8 & 9 & 0 & 0 \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & c & d \end{pmatrix}. \tag{C.13}$$

C.4.2 Example with Reference to Chap. 20

We start with

$$|h\rangle \cong \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad |v\rangle \cong \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{C.14}$$

Then it follows with $|hh\rangle \equiv |h\rangle \otimes |h\rangle$

$$|hh\rangle \cong \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ |hv\rangle \cong \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ |vh\rangle \cong \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \ |vv\rangle \cong \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

and

$$|\Phi\rangle = \frac{|hv\rangle - |vh\rangle}{\sqrt{2}} \cong \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}. \tag{C.15}$$

The measurement of the first component of this state with respect to horizontal polarization is described by

$$(|h_1\rangle \otimes I_2)(|h_1\rangle \otimes I_2)^\dagger = (|h_1\rangle \otimes I_2)(\langle h_1| \otimes I_2), \tag{C.16}$$

where I_2 is the one-operator in space 2 (for the sake of clarity, we use indices). We have

$$|h_1\rangle \otimes I_2 \cong \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \tag{C.17}$$

and it follows that

$$(|h_1\rangle \otimes I_2)(\langle h_1| \otimes I_2) \cong \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{C.18}$$

This leads to

$$(|h_1\rangle \otimes I_2)(\langle h_1| \otimes I_2)|\Phi\rangle \cong \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{C.19}$$

as is required.

The same procedure in a more compact notation reads:

$$\begin{aligned}(|h_1\rangle \otimes I_2)\,(\langle h_1| \otimes I_2)\,|\Phi\rangle &= (|h_1\rangle \otimes I_2)\,(\langle h_1| \otimes I_2)\,\frac{|h_1 v_2\rangle - |v_1 h_2\rangle}{\sqrt{2}} \\ &= (|h_1\rangle \otimes I_2)\,\frac{\langle h_1| \otimes I_2\,|h_1 v_2\rangle - \langle h_1| \otimes I_2\,|v_1 h_2\rangle}{\sqrt{2}} \\ &= (|h_1\rangle \otimes I_2)\,\frac{|v_2\rangle}{\sqrt{2}} \;=\; \frac{|h_1\rangle \otimes |v_2\rangle}{\sqrt{2}} \;=\; \frac{|h_1\rangle\,|v_2\rangle}{\sqrt{2}}.\end{aligned} \tag{C.20}$$

Appendix D
Wave Packets

D.1 General Remarks

A plane wave is not a physically realizable state: it is infinitely extended and has the same magnitude in all places and at all times (squared amplitude value). Mathematically, this is expressed by the fact that it is not square integrable. But because of the linearity of quantum mechanics, we can superpose individual waves in such a way that physically 'reasonable' expressions arise (keyword: Fourier transformation). In addition, we can construct these superpositions in such a way that they have (at least approximately) a well-defined momentum, like classical particles. We want to discuss in the following some of the characteristics of these *wave packets*.[5]

D.1.1 One-Dimensional Wave Packet

A one-dimensional wave packet generally has the form[6]:

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k)\, e^{i(kx-\omega t)} \mathrm{d}k, \tag{D.1}$$

where the amplitude function $A(k)$ is usually centered around a value K and has a pronounced maximum at K; see Fig. D.1. As an example, we can imagine a bell curve (Gaussian curve), centered at K, in which case $\psi(x,t)$ may be explicitly represented (see below). But what information can be obtained in the general case?

[5]We have already discussed some of the properties of wave packets in Chap. 15.

[6]We extract the factor $\frac{1}{\sqrt{2\pi}}$ from the integral in order to write the Fourier transform as usual.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

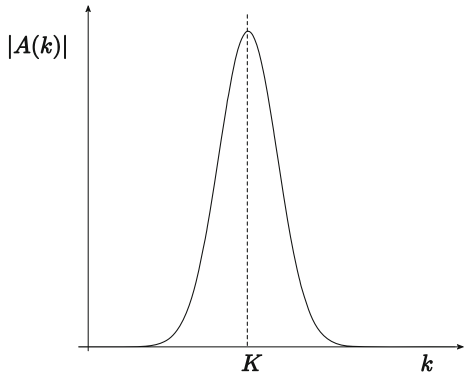


Fig. D.1 Schematic representation of the amplitude function $|A(k)|$

To answer this question, we first write $A(k) = |A(k)| e^{i\varphi(k)}$ and obtain

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |A(k)| e^{i(kx-\omega t+\varphi(k))} \mathrm{d}k = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |A(k)| e^{i\Phi(k)} \mathrm{d}k. \quad \text{(D.2)}$$

The magnitude of the integral now depends on how fast $e^{i\Phi(k)}$ oscillates in the neighborhood of K (only there is $|A(k)|$ significantly different from zero)—the faster the oscillations, the smaller becomes the integral (or its value). In general, we obtain the largest contribution if $\Phi(k)$ does not vary in the neighborhood of K, i.e. if

$$\frac{\mathrm{d}\Phi(k)}{\mathrm{d}k} = \frac{\mathrm{d}(kx - \omega t + \varphi(k))}{\mathrm{d}k} = x - \frac{\mathrm{d}\omega}{\mathrm{d}k}t + \frac{\mathrm{d}\varphi(k)}{\mathrm{d}k} = 0 \text{ for } k = K. \quad \text{(D.3)}$$

With $\omega = \frac{\hbar k^2}{2m}$, it follows that:

$$x - \frac{\hbar k}{m}t + \frac{\mathrm{d}\varphi(k)}{\mathrm{d}k} = 0; \quad \text{(D.4)}$$

or with $k = K$,

$$x = \frac{\hbar K}{m}t - \frac{\mathrm{d}\varphi(k)}{\mathrm{d}k} /_{k=K} = v_g t - \frac{\mathrm{d}\varphi(k)}{\mathrm{d}k} /_{k=K}. \quad \text{(D.5)}$$

The *group velocity* $v_g = \frac{\mathrm{d}\omega}{\mathrm{d}k}(k = K)$ denotes the propagation velocity of the wave packet, while the *phase velocity* $v_{ph} = \frac{\omega}{k}$ denotes the propagation velocity of the individual partial waves (with fixed k). Generally, it holds that $v_g \neq v_{ph}$ (the wave packet deforms in the course of time, e.g. it diverges); for $v_g = v_{ph}$, one speaks of a dispersion-free wave (e.g. electromagnetic waves in a vacuum). The concept of group velocity, moreover, makes sense only if the superposition still has a recognizable coherence, i.e. it is not fragmented.

To obtain more information about the behavior of the wave packet, we expand ω in the neighborhood of $k = K$:

$$\omega(k) = \frac{\hbar k^2}{2m} = \frac{\hbar (k - K + K)^2}{2m} = \frac{\hbar K^2}{2m} + \frac{\hbar (k - K) K}{m} + \frac{\hbar (k - K)^2}{2m} \quad \text{(D.6)}$$

and obtain with $\Omega = \frac{\hbar K^2}{2m}$ and the group velocity $v_g = \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{\hbar K}{m}$ for the phase $\Phi(k)$:

$$\begin{aligned}\Phi(k) &= kx - \omega(k)t + \varphi(k) \\ &= kx - \left[\frac{\hbar K^2}{2m} + \frac{\hbar k K}{m} - \frac{\hbar K^2}{m} + \frac{\hbar (k-K)^2}{2m}\right] t + \varphi(k) \\ &= kx + \Omega t - v_g k t - \frac{\hbar (k-K)^2}{2m} t + \varphi(k) \\ &= k\left(x - v_g t\right) + \Omega t - \frac{\hbar (k-K)^2}{2m} t + \varphi(k) .\end{aligned} \quad \text{(D.7)}$$

If we can neglect the quadratic term $(k - K)^2$ (for instance, if $c(k)$ is very narrowly concentrated about K), we obtain with (D.2) (see the exercises):

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} e^{i\Omega t} \int_{-\infty}^{\infty} A(k)\, e^{ik\left(x - v_g t\right)} \mathrm{d}k = e^{i\Omega t} \psi\left(x - v_g t, 0\right). \quad \text{(D.8)}$$

This means that, under these conditions, the wavefunction moves unchanged to the right (for $v_g > 0$) with the velocity v_g. The approximation is valid for

$$\frac{\hbar (k-K)^2}{2m} t \ll 1 \text{ or } t \ll \frac{2m}{\hbar (k-K)^2}. \quad \text{(D.9)}$$

If this inequality is satisfied, the wave packet (almost) does not disperse.

D.1.2 Example: Bell Curve

A normalized k distribution in the form of a bell curve is given by:

$$A(k) = \left(\frac{b_0^2}{\pi}\right)^{1/4} \exp\left(-\frac{b_0^2}{2}(k - K)^2\right). \quad \text{(D.10)}$$

Its maximum is at K; its width in momentum space is $\Delta k = \frac{2}{b_0}$ (the two turning points of $A(k)$ are at $k = K \pm \frac{1}{b_0}$). Thus, the corresponding wave packet is

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \left(\frac{b_0^2}{\pi}\right)^{1/4} \int_{-\infty}^{\infty} e^{-\frac{b_0^2}{2}(k-K)^2} e^{i(kx-\omega t)} \mathrm{d}k \tag{D.11}$$

with $\omega = \frac{\hbar k^2}{2m}$. This expression allows for a closed solution (calculation by variable substitution and integration in the complex plane; for details, see relevant textbooks). The result reads:

$$\psi(x,t) = \frac{1}{\sqrt{2}N(t)} \left(\frac{b_0^2}{\pi}\right)^{1/4} \exp\left(E^2(x,t) - \frac{1}{2} b_0^2 K^2\right) \tag{D.12}$$

with

$$N(t) = \sqrt{\frac{b_0^2}{2} + \frac{i\hbar t}{2m}};\ E(x,t) = \frac{Kb_0^2 + ix}{2N(t)}. \tag{D.13}$$

We are interested especially in the 'size' of $\psi(x,t)$, i.e. the squared amplitude value. From (D.12), it follows that:

$$|\psi(x,t)|^2 = \frac{1}{\sqrt{\pi}b(t)} \exp\left(-\frac{\left(x - \frac{\hbar K}{m}t\right)^2}{b^2(t)}\right) \tag{D.14}$$

with

$$b(t) = \sqrt{b_0^2 + \left(\frac{\hbar t}{mb_0}\right)^2}. \tag{D.15}$$

As we see from (D.14), $|\psi(x,t)|^2$ is very small for values that are far from $x - \frac{\hbar K}{m}t = 0$. Also from (D.14), we see directly that the group velocity is

$$v_g = \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{\hbar K}{m}. \tag{D.16}$$

This illustrates the above considerations.

By the way, the results of this section are found also in the discussion of free motion (see Chap. 5, Vol. 1).

D.1.3 Many-Dimensional Wave Packet

The generalization from one- to n-dimensional wave packets offers no fundamental surprises. We have

$$\psi(\mathbf{x}, t) = \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} A(\mathbf{k})\, e^{i(\mathbf{kx}-\omega t)} \mathrm{d}^n k, \tag{D.17}$$

where the amplitude function $c(\mathbf{k})$ is centered about a maximum at $\mathbf{k} = \mathbf{K}$. The group velocity is given by

$$\mathbf{v}_g(k) = \nabla_k \omega(k) \,/_{\mathbf{k}=\mathbf{K}}\,. \tag{D.18}$$

D.2 Potential Step and Wave Packet

As an example of an application, we consider scattering by a one-dimensional potential step. The potential is

$$V = \begin{cases} V_0 & \text{for} \quad x < 0, \quad \text{region 2} \\ 0 & \phantom{\text{for}} \quad x > 0, \quad \text{region 1}. \end{cases} \tag{D.19}$$

A quantum object is incident from the right onto the potential step; see Fig. D.2.

With the usual abbreviations

$$k^2 = \frac{2mE}{\hbar^2};\ \kappa^2 = \frac{2m}{\hbar^2}(V_0 - E)\ \text{for}\ V_0 > E;\ k'^2 = \frac{2m}{\hbar^2}(E - V_0)\ \text{for}\ V_0 < E \tag{D.20}$$

and

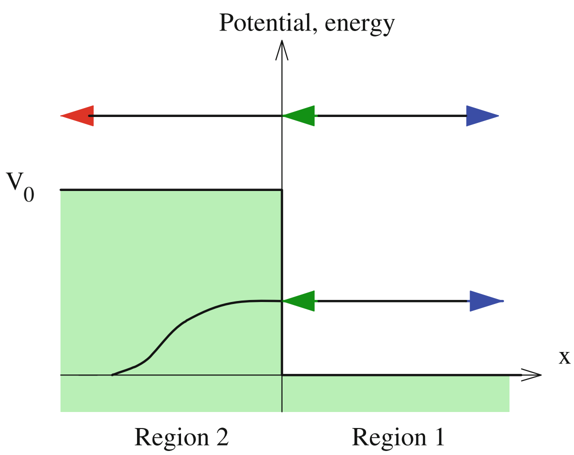


Fig. D.2 Situation at the potential step for $E > V_0$ (*above*) and $E < V_0$ (*below*). The *horizontal lines* indicate a wave, the *curved line* an exponential decay

$$\gamma(k) = \begin{cases} \kappa = \sqrt{k_0^2 - k^2} \\ -ik' = -i\sqrt{k^2 - k_0^2} \end{cases} \text{for } \begin{matrix} E < V_0 \\ E > V_0 \end{matrix} \text{ or } \begin{matrix} k < k_0 \\ k > k_0 \end{matrix};\ k_0 = \sqrt{\frac{2m}{\hbar^2} V_0}, \tag{D.21}$$

a partial solution for fixed k, as shown in Chap. 15, is given by:

$$\varphi_1 = e^{-ikx} + \frac{ik + \gamma}{ik - \gamma} e^{ikx};\ \varphi_2 = \frac{2ik}{ik - \gamma} e^{\gamma x} \tag{D.22}$$

with $k > 0$, $\kappa > 0$.

From this, we obtain the total solution by integration over the continuous index $k > 0$:

$$\begin{aligned} \Psi_1(x,t) &= \int_0^\infty c(k) \left(e^{-ikx} + \tfrac{ik+\gamma}{ik-\gamma} e^{ikx} \right) e^{-i\omega t} \mathrm{d}k;\ x > 0 \\ \Psi_2(x,t) &= \int_0^\infty c(k) \tfrac{2ik}{ik-\gamma} e^{\gamma x} e^{-i\omega t} \mathrm{d}k;\ x < 0 \end{aligned} \quad \text{with } \omega = \frac{\hbar k^2}{2m}. \tag{D.23}$$

$c(k)$ is a function of k which, as discussed above, is nonzero only in a neighborhood of a certain momentum K (as e.g. in the bell curve).

With the definition of γ, it follows that

$$\begin{aligned} \Psi_1(x,t) &= \int_0^\infty c(k) e^{-i(kx+\omega t)} \mathrm{d}k - \int_0^{k_0} c(k) e^{i(kx-\omega t) + 2i \arctan k/\kappa} \mathrm{d}k \\ &+ \int_{k_0}^\infty c(k) \tfrac{k-k'}{k+k'} e^{i(kx-\omega t)} \mathrm{d}k \\ \Psi_2(x,t) &= \int_0^{k_0} c(k) \tfrac{2ik}{ik-\kappa} e^{\kappa x} e^{-i\omega t} \mathrm{d}k + \int_{k_0}^\infty c(k) \tfrac{2k}{k+k'} e^{-i(k'x+\omega t)} \mathrm{d}k. \end{aligned} \tag{D.24}$$

We see that $\Psi_1(x,t)$ contains two types of waves: On the one hand, there are waves travelling from right to left which in our model concept represent the incoming quantum object ($\varphi_{\text{ein}} \sim e^{-i(kx+\omega t)}$); on the other hand there are waves travelling from left to right that represent the reflected quantum object ($\varphi_{\text{refl}} \sim e^{i(kx-\omega t)}$). $\Psi_2(x,t)$ for $k < k_0$ is the exponentially-damped term ($\sim e^{\kappa x}$), and for $k > k_0$, it is the transmitted part of the wavefunction ($\varphi_{\text{trans}} \sim e^{-i(k'x+\omega t)}$).[7]

Depending on the choice of $c(k)$, one can create very complicated wavefunctions. To allow the comparison with classical behavior, we again choose $c(k)$ in such a way that it is centered around a value K (i.e. we have $\Delta k \ll K$) and has a pronounced maximum there. Furthermore, we restrict the discussion to two cases: (1)

[7] We recall that the wavefunction does not describe the object itself, but allows for the calculation of its position probability.

The maximum of $c(k)$ is at $K < k_0$; outside of the interval $(0, k_0)$, $c(k)$ vanishes. (2) The maximum of $c(k)$ is at $K > k_0$; outside of the interval (k_0, ∞), $c(k)$ vanishes.

Case 1: For the energy of all the partial waves (and therefore for the total wave), we take $E < V_0$. From the point of view of classical physics, this case corresponds to an object which is incident first from right to left with velocity v at the potential step, then is reflected and travels back at the same velocity from left to right. The quantum-mechanical behavior differs from this in two (causally related) points: First, the wavefunction penetrates into the classically forbidden region 2 (i.e. $x < 0$); on the other hand (and as a result of this intrusion), the reflected wave experiences a phase delay.

To see this more closely, we consider the wavefunction. Because of $c(k) = 0$ for $k > k_0$, we have:

$$\Psi_1(x,t) = \int_0^\infty c(k)\, e^{-i(kx+\omega t)} \mathrm{d}k - \int_0^{k_0} c(k)\, e^{i(kx-\omega t)+2i \arctan k/\kappa} \mathrm{d}k$$

$$\Psi_2(x,t) = \int_0^{k_0} c(k)\, \frac{2ik}{ik-\kappa} e^{\kappa x} e^{-i\omega t} \mathrm{d}k. \tag{D.25}$$

Incoming ($\sim e^{-i(kx+\omega t)}$) and reflected ($\sim e^{i(kx-\omega t)}$) wave components have the same absolute value of their amplitudes; the group velocity for both components of the wave packet is $\left|v_g\right| = \frac{\mathrm{d}\omega}{\mathrm{d}k} /_{k=K} = \frac{\hbar K}{m}$. Thus far, the quantum-mechanical behavior corresponds to the classical solution. The main difference lies in the fact that the wavefunction does not vanish identically for $x < 0$. Intuitively, this means that the quantum object penetrates into the classically forbidden region (essentially the beginning of the tunnel effect). This leads to a (k-dependent) phase shift τ of the reflected partial wave:

$$\tau = \frac{2}{\omega} \arctan \frac{k}{\kappa} = \frac{2m}{\hbar k} \arctan \frac{k}{\kappa}. \tag{D.26}$$

Case 2: For the energy of all the partial waves (and therefore the total wave), we take $E > V_0$. From the point of view of classical physics, this case corresponds to an object which is first incident from right to left with velocity v onto the potential step, and from there continues to travel in the same direction, but at a lower speed, v'. The quantum-mechanical behavior differs in one respect: it includes a reflection at the step.

We can see this directly from the wavefunction:

$$\Psi_1(x,t) = \int_0^\infty c(k)\, e^{-i(kx+\omega t)} \mathrm{d}k + \int_{k_0}^\infty c(k)\, \frac{k-k'}{k+k'} e^{i(kx-\omega t)} \mathrm{d}k$$

$$\Psi_2(x,t) = \int_{k_0}^{\infty} c(k) \frac{2k}{k + k'} e^{-i(k'x + \omega t)} dk. \tag{D.27}$$

The amplitudes of the reflected wave components vanish with increasing energy E, proportionally to $\frac{V_0}{4E}$. The reflection takes place (in contrast to case 1) with no phase delay, i.e. instantaneously. The group velocity of the transmitted component is obtained from the stationarity of the total phase:

$$0 = \frac{d(k'x + \omega t)}{dk} = \frac{k}{k'}x + \frac{\hbar k}{m}t = \frac{k}{k'}\left(x + \frac{\hbar k'}{m}t\right), \tag{D.28}$$

giving $v_g' = \frac{\hbar}{m}\sqrt{K^2 - k_0^2}$, compared to $v_g = \frac{\hbar K}{m}$ for the incoming component. Hence, as in the classical case, the group velocity of the transmitted component is lower as that of the incoming component.

We note that the wavefunction *always* has a reflected part under these circumstances (albeit possibly a very small one). Intuitively speaking, quantum mechanics thus allows for the scattering of a cannon ball by a snowflake (as the historical example goes). This contradicts our everyday experience, but it seems less strange when one thinks not of matter, but of light. We consider the propagation of light in a non-absorbing medium with a variable refractive index. In case 1, we have a change from a real refractive index (region 1) to an imaginary one (region 2) and correspondingly total reflection. In case 2, we have a sudden change of the value of the real refractive index, which always causes a partial reflection of the light.

In determining the group velocity, we have tacitly assumed that the maxima of the functions $F(k) = \frac{k-k'}{k+k'}c(k)$ and $G(k) = \frac{2k}{k+k'}c(k)$ are located approximately at $k = K$. We want to check this briefly. With

$$\frac{2k}{k + k'} = \frac{k - k'}{k + k'} + 1;\ k' = \sqrt{k^2 - k_0^2};\ \frac{dk'}{dk} = \frac{k}{k'};\ \frac{d}{dk}\frac{2k}{k + k'} = \frac{2}{k'}\frac{k' - k}{k + k'}, \tag{D.29}$$

we obtain for the position of the maxima the conditional equations

$$\begin{aligned} \frac{d}{dk}c(k)\frac{k - k'}{k + k'} &= c^{(1)}\frac{k - k'}{k + k'} + c\frac{2}{k'}\frac{k' - k}{k + k'} = 0 \\ \frac{d}{dk}c(k)\frac{2k}{k + k'} &= c^{(1)}\frac{2k}{k + k'} + c\frac{2}{k'}\frac{k' - k}{k + k'} = 0 \end{aligned} \tag{D.30}$$

or

$$k'c^{(1)} - 2c = 0;\ k'kc^{(1)} + c(k' - k) = 0. \tag{D.31}$$

As a typical distribution, we insert a bell curve (D.10) for $c(k)$; with $c^{(1)} = -b_0^2(k - K)c$, it follows that

$$k'b_0^2\,(k-K)\;+\;2=0;\;\;k'kb_0^2\,(k-K)\;+\;k-k'=0. \tag{D.32}$$

Instead of looking for exact (as possible) solutions of these equations, we use approximations according to our more qualitative approach. We use the fact that the width of the distribution in momentum space is given by $\Delta k = \frac{2}{b_0}$. Thus, the last two equations can be written as

$$k-K=-\frac{2}{k'b_0^2}=-\frac{1}{2k'}\,(\Delta k)^2\,;\;\;k-K=\frac{k'-k}{k'kb_0^2}=-\frac{k-k'}{4k'k}\,(\Delta k)^2 \tag{D.33}$$

or

$$k=K\left[1-\frac{(\Delta k)^2}{2k'K}\right];\;\;k=K\left[1-\frac{(\Delta k)^2}{2k'K}\frac{k-k'}{2k}\right]. \tag{D.34}$$

Due to $\Delta k \ll K$ and $k_0 < k$, we can replace k approximately by K and k' by $K'=\sqrt{K^2-k_0^2}$, obtaining for the position of the maxima:

$$k=K\left[1-\frac{(\Delta k)^2}{2K'K}\right];\;\;k=K\left[1-\frac{(\Delta K)^2}{2K'K}\frac{K-K'}{2K}\right]. \tag{D.35}$$

We see directly that for sufficiently narrow distributions (essentially $(\Delta k)^2 \ll K'K$), the maxima of the two distributions are at $k \approx K$, and thus our discussion concerning the group velocity was consistent.

D.3 Exercises

1. The function $\psi\,(x,t)$ is given as

$$\psi\,(x,t)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}A\,(k)\,e^{ik\left(x-v_g t\right)}\mathrm{d}k. \tag{D.36}$$

Show that:

$$\psi\,(x,t)=\psi\left(x-v_g t,0\right). \tag{D.37}$$

Solution: We have

$$\psi\,(x,0)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}A\,(k)\,e^{ikx}\mathrm{d}k. \tag{D.38}$$

It follows that

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k)\, e^{ik\left(x - v_g t\right)} \mathrm{d}k \underset{y = x - v_g t}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k)\, e^{iky} \mathrm{d}k$$
$$= \psi(y, 0) = \psi\left(x - v_g t, 0\right). \tag{D.39}$$

2. For which times can one neglect the broadening of the Gaussian wave packet (D.12)?
Solution: The width of the distribution in position space ($\Delta x = \sqrt{2} b(t)$) is given by (D.15), i.e. by

$$b(t) = \sqrt{b_0^2 + \left(\frac{\hbar t}{m b_0}\right)^2} \approx b_0 \left[1 + \frac{1}{2}\left(\frac{\hbar t}{m b_0^2}\right)^2\right]. \tag{D.40}$$

The packet has practically not broadened for times

$$t \ll \sqrt{2} \frac{m b_0^2}{\hbar}. \tag{D.41}$$

Because of $\Delta k = \frac{2}{b_0}$, it follows that

$$t \ll \frac{4\sqrt{2} m}{\hbar (\Delta k)^2} \tag{D.42}$$

in accordance with (D.9).
3. The relativistic energy-momentum relation is given by

$$E^2 = m_0^2 c^4 + c^2 p^2. \tag{D.43}$$

Determine the group velocity and the phase velocity v_g and v_{ph}. Show that $v_g v_{ph} = c^2$. Which velocity is greater than c?
Solution: We have:

$$v_g = \frac{\mathrm{d}E}{\mathrm{d}p} = \frac{c^2 p}{E};\ v_{ph} = \frac{E}{p}. \tag{D.44}$$

For the product, it follows immediately that:

$$v_g v_{ph} = \frac{c^2 p}{E} \frac{E}{p} = c^2. \tag{D.45}$$

For the phase velocity, we find

$$v_{ph}^2 = \frac{m_0^2 c^4 + c^2 p^2}{p^2} = c^2 + \frac{m_0^2 c^4}{p^2} \geq c^2. \tag{D.46}$$

Appendix E
Laboratory System, Center-of-Mass System

The hydrogen atom is a concrete example of the general case of a two-body problem, where two interacting masses or quantum objects are considered without external forces. The total energy of this system is composed of the kinetic energies of the two bodies and the potential energy, i.e. the interaction energy V between them:

$$E = E_{\text{kin}} + E_{\text{pot}} = \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} + V. \tag{E.1}$$

We assume in the following that the potential depends only on the relative distance $\mathbf{r}_1 - \mathbf{r}_2$, i.e. on $V = V(\mathbf{r}_1 - \mathbf{r}_2)$. Under this assumption, one can reduce the problem to an equivalent one-body problem; specialization to the Coulomb interaction of point charges leads in the quantum-mechanical treatment to the well-known form of the hydrogen spectrum.

E.1 The Equivalent One-Body Problem

For a simpler description of the problem, we introduce new coordinates, namely *center-of-mass coordinates* and *relative coordinates*:

$$\begin{aligned} \mathbf{R} &= \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} \text{ and } \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \\ \mathbf{R} &= (X, Y, Z) \text{ and } \mathbf{r} = (x, y, z) \end{aligned} \tag{E.2}$$

as well as the *total mass* and the *reduced mass*:

$$M = m_1 + m_2 \text{ and } \mu = \frac{m_1 m_2}{m_1 + m_2}. \tag{E.3}$$

One can perform this transformation classically and then change to quantum mechanics, or proceed *vice versa*. In the following, both approaches are discussed.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

E.2 Transformation Laboratory System → Center-of-Mass System

E.2.1 First Transformation, Then Transition to Quantum Mechanics

The inverse transformations to (E.2) read:

$$\mathbf{r}_1 = \mathbf{R} + \frac{m_2}{M}\mathbf{r} \text{ and } \mathbf{r}_2 = \mathbf{R} - \frac{m_1}{M}\mathbf{r}. \tag{E.4}$$

With $\mathbf{p}_1 = m_1\dot{\mathbf{r}}_1$ etc., taking derivatives with respect to time leads to:

$$\mathbf{p}_1 = m_1\dot{\mathbf{R}} + \mu\dot{\mathbf{r}} \text{ and } \mathbf{p}_2 = m_2\dot{\mathbf{R}} - \mu\dot{\mathbf{r}}. \tag{E.5}$$

For the kinetic energy, we find:

$$E_{\text{kin}} = \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}\mu\dot{\mathbf{r}}^2 = \frac{\mathbf{P}^2}{2M} + \frac{\mathbf{p}^2}{2\mu} \tag{E.6}$$

with the *center-of-mass momentum* and the *relative momentum*

$$\mathbf{P} = M\dot{\mathbf{R}} \text{ and } \mathbf{p} = \mu\dot{\mathbf{r}} \tag{E.7}$$

The total energy is thus

$$E = E_{\text{kin}} + E_{\text{pot}} = \frac{\mathbf{P}^2}{2M} + \frac{\mathbf{p}^2}{2\mu} + V(\mathbf{r}). \tag{E.8}$$

We now go into the *center-of-mass system*, where $\mathbf{P} = 0$; it follows that:

$$E = \frac{\mathbf{p}^2}{2\mu} + V(\mathbf{r}) \text{ in the center-of-mass system.} \tag{E.9}$$

This problem depends only on the relative coordinate; it is called the (classical) one-body problem. If we now proceed to quantum mechanics, we obtain from the last equation in the usual way, i.e. setting $\mathbf{p} = \frac{\hbar}{i}\nabla$, the Hamiltonian of the relative motion

$$H = -\frac{\hbar^2}{2\mu}\nabla^2 + V(\mathbf{r}). \tag{E.10}$$

E.2.2 First Transition to Quantum Mechanics, Then Transformation

We start with (E.1) and obtain

$$H = -\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(\mathbf{r}_1 - \mathbf{r}_2). \tag{E.11}$$

Now we have to rearrange the nabla operators ∇_1 and ∇_2 using

$$\nabla_n = \left(\frac{\partial}{\partial x_n}, \frac{\partial}{\partial y_n}, \frac{\partial}{\partial z_n}\right) \tag{E.12}$$

by means of the variable transformation (E.2) to give the nabla operators $\nabla_{\mathbf{R}}$ and $\nabla_{\mathbf{r}}$. We have (chain rule)

$$\frac{\partial}{\partial x_1} = \frac{\partial X}{\partial x_1}\frac{\partial}{\partial X} + \frac{\partial Y}{\partial x_1}\frac{\partial}{\partial Y} + \frac{\partial Z}{\partial x_1}\frac{\partial}{\partial Z} + \frac{\partial x}{\partial x_1}\frac{\partial}{\partial x} + \frac{\partial y}{\partial x_1}\frac{\partial}{\partial y} + \frac{\partial z}{\partial x_1}\frac{\partial}{\partial z}. \tag{E.13}$$

With

$$\frac{\partial X}{\partial x_1} = \frac{m_1}{M} \text{ and } \frac{\partial x}{\partial x_1} = 1, \tag{E.14}$$

it follows that

$$\frac{\partial}{\partial x_1} = \frac{m_1}{M}\frac{\partial}{\partial X} + \frac{\partial}{\partial x}, \tag{E.15}$$

and similarly for the other expressions.[8] Overall, we find:

$$\nabla_1 = \nabla_{\mathbf{r}} + \frac{m_1}{M}\nabla_{\mathbf{R}} \text{ and } \nabla_2 = -\nabla_{\mathbf{r}} + \frac{m_2}{M}\nabla_{\mathbf{R}}.$$

Inserting into the Hamiltonian (E.11) yields:

$$H = -\frac{\hbar^2}{2M}\nabla_{\mathbf{R}}^2 - \frac{\hbar^2}{2\mu}\nabla_{\mathbf{r}}^2 + V(\mathbf{r}). \tag{E.16}$$

The time-dependent Schrödinger equation

$$i\hbar\dot{\Psi} = H\Psi \text{ with } \Psi = \Psi(\mathbf{R}, \mathbf{r}, t) \tag{E.17}$$

---

[8] One can write this and similar conversions compactly using the (mathematically very sloppy) notation $\nabla_1 = \frac{\partial}{\partial \mathbf{r}_1}$ (this does not mean that one 'divides' by the vector $\mathbf{r}_1$; it is only a different notation for the nabla operator). Then it follows e.g. that $\frac{\partial}{\partial \mathbf{r}_1} = \frac{\partial \mathbf{r}}{\partial \mathbf{r}_1}\frac{\partial}{\partial \mathbf{r}} + \frac{\partial \mathbf{R}}{\partial \mathbf{r}_1}\frac{\partial}{\partial \mathbf{R}}$, etc.; transformations of this kind can thus be performed quite cleverly and with little paperwork. But one has to know what this notation means and how to deal with it.

yields with the usual separation *ansatz*

$$\Psi(\mathbf{R}, \mathbf{r}, t) = e^{-i\omega t}\Phi(\mathbf{R}, \mathbf{r}) \text{ with } E_{\text{total}} = \hbar\omega \tag{E.18}$$

the stationary total Schrödinger equation

$$E_{\text{total}}\Phi(\mathbf{R}, \mathbf{r}) = H\Phi(\mathbf{R}, \mathbf{r}) = \left(-\frac{\hbar^2}{2M}\nabla_{\mathbf{R}}^2 - \frac{\hbar^2}{2\mu}\nabla_{\mathbf{r}}^2 + V(\mathbf{r})\right)\Phi(\mathbf{R}, \mathbf{r}). \tag{E.19}$$

The coordinate **R** appears only in the first term on the right-hand side; thus, a separation *ansatz* makes sense:

$$\Phi(\mathbf{R}, \mathbf{r}) = g(\mathbf{R})\psi(\mathbf{r}). \tag{E.20}$$

Inserting and applying the known argumentation gives with

$$E_{\text{total}} = E_{\mathbf{R}} + E_{\mathbf{r}} \tag{E.21}$$

finally a split of the equation into two equations, namely:

1. an equation for the center-of-mass, in fact a free motion:

$$-\frac{\hbar^2}{2M}\nabla_{\mathbf{R}}^2 g(\mathbf{R}) = E_{\mathbf{R}} g(\mathbf{R}); \tag{E.22}$$

2. an equation for the relative motion (which contains the interaction between the two quantum objects):

$$-\frac{\hbar^2}{2\mu}\nabla_{\mathbf{r}}^2\psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) = E_{\mathbf{r}}\psi(\mathbf{r}). \tag{E.23}$$

Thus we have again (E.10). In other words, no matter whether we proceed first classically or first quantum mechanically, we end up (as we must) with the same result, i.e. (E.23). This is the *equivalent one-body problem*. It differs basically only in one point from the problem of a quantum object of mass m in a potential V, namely in the occurrence of the reduced mass μ instead of the mass m.

Appendix F
Analytic Treatment of the Hydrogen Atom

In this section, we consider the detailed derivation of the solution of the radial equation of the hydrogen atom by a power series approach.

F.1 Nonrelativistic Case: Schrödinger equation

We start from the radial equation in the form (17.25), namely:

$$\frac{\mathrm{d}^2 u_{nl}(r)}{\mathrm{d}r^2} - \left(\frac{l(l+1)}{r^2} - \frac{2\mu\gamma}{\hbar^2 r} + \frac{2\mu}{\hbar^2}\left|E_n\right|\right) u_{nl}(r) = 0. \tag{F.1}$$

For the meaning of the individual quantities, see Chap. 17

Simplification of the Constants

In (F.1), the appearance of five constants (l, μ, γ, $\hbar$, and E_n) is annoying. For simplicity, we transform to a new variable ρ:

$$u_{nl}(r) = S_{nl}(\xi r) = S_{nl}(\rho) \tag{F.2}$$

where we choose ρ or ξ suitably. Inserting in the radial equation yields:

$$\frac{\mathrm{d}^2 S_{nl}(\rho)}{\mathrm{d}\rho^2} - \left(\frac{l(l+1)}{\rho^2} - \frac{2\mu\gamma}{\hbar^2 \xi\rho} + \frac{2\mu}{\hbar^2\xi^2}\left|E_n\right|\right) S_{nl}(\rho) = 0. \tag{F.3}$$

We choose

$$\xi = \sqrt{\frac{8\mu\left|E_n\right|}{\hbar^2}} \tag{F.4}$$

and obtain with the abbreviation

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

$$c = \gamma\sqrt{\frac{\mu}{2\hbar^2 |E_n|}} \tag{F.5}$$

the following equation for $S_{nl}(\rho)$, with only two constants:

$$\frac{d^2 S_{nl}(\rho)}{d\rho^2} - \left(\frac{l(l+1)}{\rho^2} - \frac{c}{\rho} + \frac{1}{4}\right) S_{nl}(\rho) = 0. \tag{F.6}$$

Separating the Behavior for $r \to 0, r \to \infty$

For $\rho \to \infty$ (i.e. $r \to \infty$), this equation becomes

$$\frac{d^2 S_{nl}(\rho)}{d\rho^2} - \frac{1}{4} S_{nl}(\rho) = 0 \tag{F.7}$$

with the solution[9]:

$$S_{nl}(\rho) = e^{-\rho/2}, \tag{F.8}$$

and for $\rho \to 0$, we obtain approximately

$$\frac{d^2 S_{nl}(\rho)}{d\rho^2} - \frac{l(l+1)}{\rho^2} S_{nl}(\rho) = 0 \tag{F.9}$$

with the solutions

$$S_{nl}(\rho) = \begin{cases} \rho^{l+1}\text{: regular solution} \\ \rho^{-l}\text{: irregular solution} \end{cases} \tag{F.10}$$

The irregular solution has a pole at zero; we exclude it as unphysical. We consider the behavior for $r \to 0$ and for $r \to \infty$ and obtain as our *ansatz*:

$$S_{nl}(\rho) = \rho^{l+1} e^{-\rho/2} f_{nl}(\rho). \tag{F.11}$$

Inserting this into (F.6) yields

$$\rho \frac{d^2 f_{nl}(\rho)}{d\rho^2} + [2(l+1) - \rho] \frac{d f_{nl}(\rho)}{d\rho} + [c - l - 1] f_{nl}(\rho) = 0. \tag{F.12}$$

Solution by Power Series

To solve this differential equation, we use a power series *ansatz* of the form

---

[9]We note that the exact solution of an asymptotically-approximated differential equation need not be identical with the asymptotic solution of the exact differential equation. However, that does not matter here, as we are seeking only a clever *ansatz* to simplify the problem, but not (at this point) the exact solution.

$$f_{nl}(\rho) = \sum_{k=0}^{\infty} a_k \rho^k. \tag{F.13}$$

Inserting gives[10]:

$$\begin{aligned} 0 = &\sum_{k=0}^{\infty} k(k+1)a_{k+1}\rho^k + 2(l+1)\sum_{k=0}^{\infty}(k+1)a_{k+1}\rho^k \\ &- \sum_{k=0}^{\infty} k a_k \rho^k + [c - l - 1]\sum_{k=0}^{\infty} a_k \rho^k. \end{aligned} \tag{F.14}$$

Comparing the coefficients of like powers leads to

$$k(k+1)a_{k+1} + 2(l+1)(k+1)a_{k+1} - k a_k + [c - l - 1]\, a_k = 0 \tag{F.15}$$

or

$$a_{k+1} = \frac{k + l + 1 - c}{(k+1)\,(k + 2l + 2)} a_k. \tag{F.16}$$

Using this relation, we can thus recursively compute all the coefficients of the power series, if we specify a_0 (the zero-th term is determined only by the normalization of the total wavefunction). However, it is still unclear what the radius of convergence of the power series (F.13) is and whether the solution is physically acceptable. To answer these questions, we first note that

$$\frac{a_{k+1}}{a_k} \underset{k\to\infty}{\to} \frac{1}{k+1}. \tag{F.17}$$

By the ratio test, the series converges (because of $\lim\limits_{k\to\infty} \frac{a_{k+1}}{a_k} = \lim\limits_{k\to\infty} \frac{1}{k+1} = 0 < 1$). That is, the power series has the same convergence behavior as e^{ρ} (because of $e^{\rho} = \sum c_k \rho^k = \sum \rho^k / k!$ and $c_{k+1}/c_k = 1/(k+1)$). However, although the power series converges (the radius of convergence is in fact ∞), it is not acceptable for physical reasons. This is because, as we can see from (F.11), the asymptotic behavior of the radial function $R_{nl} = u_{nl(r)}/r$ would be given in this case by $\rho^l e^{\rho/2}$ and, accordingly, R_{nl} would not be square integrable.

We examine this in more detail. In addition to the function

$$f_{nl}(\rho) = \sum_{k=0}^{\infty} a_k \rho^k, \tag{F.18}$$

---

[10]Here, we use rearrangements such as

$$\sum_{k=0}^{\infty} k a_k \rho^{k-1} = \sum_{k=1}^{\infty} k a_k \rho^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} \rho^k.$$

we consider the 'comparison function'

$$e^{\lambda\rho} = \sum_{k=0}^{\infty} b_k \rho^k = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \rho^k \text{ with } 0 < \lambda < 1. \tag{F.19}$$

Now there is a K such that for $k \geq K$, we have:

$$\frac{a_{k+1}}{a_k} > \frac{b_{k+1}}{b_k} > 0. \tag{F.20}$$

One can easily see why, since due to (F.16),

$$\frac{a_{k+1}}{a_k} = \frac{k+l+1-c}{(k+1)(k+2l+2)} > \frac{b_{k+1}}{b_k} = \frac{\lambda}{k+1} > 0. \tag{F.21}$$

Solving this equation for k yields

$$K > \frac{(l+1)(2\lambda - 1) + c}{1-\lambda} > c - l - 1. \tag{F.22}$$

Clearly, this condition is the most stringent (K is largest) if we restrict ourselves to the range $1/2 < \lambda < 1$; this we take to be the case from now on. We split:

$$\begin{aligned} f_{nl}(\rho) &= \sum_{k=0}^{K-1} a_k \rho^k f(\rho) + \sum_{k=K}^{\infty} a_k \rho^k = P(\rho) + \sum_{k=K}^{\infty} a_k \rho^k \\ &= P(\rho) + \sum_{m=0}^{\infty} a_{K+m} \rho^{K+m} \end{aligned} \tag{F.23}$$

and

$$e^{\lambda\rho} = \sum_{k=0}^{K-1} b_k \rho^k + \sum_{k=K}^{\infty} b_k \rho^k = Q(\rho) + \sum_{k=K}^{\infty} b_k \rho^k = Q(\rho) + \sum_{m=0}^{\infty} b_{K+m} \rho^{K+m}, \tag{F.24}$$

and note that:

$$a_{K+m} = a_K \prod_{l=0}^{m-1} \frac{a_{K+l+1}}{a_{K+l}} > a_K \prod_{l=0}^{m-1} \frac{b_{K+l+1}}{b_{K+l}} = a_K \frac{b_{K+m}}{b_K}. \tag{F.25}$$

Now we can estimate as follows:

$$f_{nl}(\rho) - P(\rho) = \sum_{m=0}^{\infty} a_{K+m} \rho^{K+m} > \frac{a_K}{b_K} \sum_{m=0}^{\infty} b_{K+m} \rho^{K+m} = \frac{a_K}{b_K} [e^{\lambda\rho} - Q(\rho)]. \tag{F.26}$$

This inequality holds for all ρ, i.e. also for $\rho \to \infty$. But in that case, we can neglect the polynomials P and Q in comparison with the corresponding function, so that we obtain:

$$f_{nl}(\rho) \geq \frac{a_K}{b_K} e^{\lambda\rho}, \tag{F.27}$$

and we have finally

$$S_{nl}(\rho) = \rho^{l+1} e^{-\rho/2} f_{nl}(\rho) \geq \frac{a_K}{b_K} \rho^{l+1} e^{(\lambda - 1/2)\rho}. \tag{F.28}$$

Because of $1/2 < \lambda < 1$, we have $R_{nl} = u_{nl(r)}/r \underset{r\to\infty}{\to} \infty$, and therefore the radial function would not be square integrable and would not be physically meaningful. In other words, the power series (F.18) always gives a physically meaningless result.

There is only one way out of this situation, namely if $f_{nl}(\rho)$ in (F.13) is *not* an infinite power series, but a rather polynomial.[11] For some natural number m, it must therefore hold that $a_m = 0$, because then the radial function behaves for large r essentially as (a polynomial in r times $e^{-\rho/2}$), and thus is square integrable. So we have to require that the numerator in (F.16) vanishes—in other words, it must hold that $c \in \mathbb{N}$. This is exactly why we rename c to n and call this number the *principal quantum number*. In addition, from (F.16), it follows for $l = n - 1 - k$ that for a given n, the quantum number l can have only the values

$$l = 0, 1, \ldots, n-2, n-1. \tag{F.29}$$

With (F.5), we thus have the identity:

$$n = \gamma \sqrt{\frac{\mu}{2\hbar^2 |E_n|}}; n \in \mathbb{N}, \tag{F.30}$$

and solving for $|E_n|$ yields

$$|E_n| = \frac{\mu\gamma^2}{2\hbar^2} \frac{1}{n^2}. \tag{F.31}$$

Hence, the energy spectrum is discrete for negative energies E, i.e. for bound states.

F.2 Relativistic Case: Dirac equation

The Dirac equation describes the Hydrogen spectrum with considerably more precision than the Schrödinger equation. Historically, the good agreement (fine structure etc.) with the experimental results was an important contribution to the triumph of the Dirac equation and the underlying ideas.

[11] As it turns out, these polynomials are the associated Laguerre polynomials; see Chap. 17.

In the following, we will sketch the proceeding in an abbreviated manner, leaving some steps to the reader. We refer to Sect. 16.5 (addition of angular momenta), Chap. 17 (Hydrogen atom) and Sect. 19.3 (Hydrogen: fine structure). Note that in this section we write m_0 for the mass of the electron in order to avoid confusion with the z-component m of the angular momentum.

F.2.1 From 4-Spinor to 2-Spinors

We start with the Dirac equation in the form

$$i\hbar\frac{\partial}{\partial t}\psi = c\boldsymbol{\alpha}\left(\mathbf{p}-q\mathbf{A}\right)\psi + q\Phi\psi + m_0c^2\beta\psi \tag{F.32}$$

where m_0 is the rest mass of the electron. Considering an electron in a Coulomb potential, we have

$$\mathbf{A} = 0\ ;\ q\Phi = -\frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r} = V\left(r\right) \tag{F.33}$$

where Z is the proton number and e the charge of the electron. The Hamilton operator reads

$$H = c\boldsymbol{\alpha}\mathbf{p}+V\left(r\right) + m_0c^2\beta. \tag{F.34}$$

The method to calculate the eigenvalues is in parts similar to the non-relativistic case though there are fundamental differences. For instance, the nonrelativistic state is 1-dimensional, whereas in the Dirac case, the state is a 4-spinor. In addition, in the Schrödinger case as considered in Chap. 17, we have the orbital angular momentum $\mathbf{l}$ only, whereas here we have to take into account the spin $\mathbf{s}$ in addition. The total angular momentum is given by $\mathbf{j} = \mathbf{l} + \mathbf{s}$.

Since we consider a central field, the total angular momentum $\mathbf{j}$ commutes with the Hamiltonian (F.34), so we can construct eigenfunctions of simultaneously H, j^2 and j_z. In addition, the Hamiltonian (F.34) is invariant against space reflection given by (see Appendix U, Vol. 1).

$$P = \beta P^{(\mathbf{x})}\ ;\ P^{(\mathbf{x})} : \mathbf{x} \to -\mathbf{x}. \tag{F.35}$$

We know that P commutes with H, that $P^2 = 1$ and that, correspondingly, P has the eigenvalues ± 1. In other words, H allows for eigenfunctions with defined parity, even or odd.

Let us go now into details. We are searching solutions for the stationary equation with $E > 0$:

$$E\Psi = c\ \boldsymbol{\alpha}\cdot\boldsymbol{p}\boldsymbol{\Psi}+ V\left(r\right)\Psi + m_0c^2\beta\Psi. \tag{F.36}$$

Since the problem has spherical symmetry, a description in terms of spherical coordinates will be favourable. In view of the block structure of the matrices $\boldsymbol{\alpha}$ and β, we write the 4-spinor Ψ as

$$\Psi = \begin{pmatrix} \Phi \\ X \end{pmatrix} \tag{F.37}$$

where Φ and X are 2-spinors. Thus, in the standard representation of $\boldsymbol{\alpha}$ and β follows

$$E\begin{pmatrix} \Phi \\ X \end{pmatrix} = c\begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}\mathbf{p}\begin{pmatrix} \Phi \\ X \end{pmatrix} + V(r)\begin{pmatrix} \Phi \\ X \end{pmatrix} + m_0c^2\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \Phi \\ X \end{pmatrix} \tag{F.38}$$

or

$$\begin{aligned} E\Phi &= c\,\boldsymbol{\sigma}\cdot\mathbf{p}X + V(r)\,\Phi + m_0c^2\Phi \\ EX &= c\,\boldsymbol{\sigma}\cdot\mathbf{p}\Phi + V(r)\,X - m_0c^2X \end{aligned}. \tag{F.39}$$

In order that the 4-spinor Ψ has a defined parity, the 2-spinors have to fulfill

$$\beta P_0\begin{pmatrix} \Phi(\mathbf{r}) \\ X(\mathbf{r}) \end{pmatrix} = \beta\begin{pmatrix} \Phi(-\mathbf{r}) \\ X(-\mathbf{r}) \end{pmatrix} = \begin{pmatrix} \Phi(-\mathbf{r}) \\ -X(-\mathbf{r}) \end{pmatrix} \overset{!}{=} \pm\begin{pmatrix} \Phi(\mathbf{r}) \\ X(\mathbf{r}) \end{pmatrix} \tag{F.40}$$

or explicitly

$$\begin{pmatrix} \Phi(-\mathbf{r}) \\ X(-\mathbf{r}) \end{pmatrix} \overset{!}{=} \begin{pmatrix} \Phi(\mathbf{r}) \\ -X(\mathbf{r}) \end{pmatrix} \text{ or } \begin{pmatrix} \Phi(-\mathbf{r}) \\ X(-\mathbf{r}) \end{pmatrix} \overset{!}{=} \begin{pmatrix} -\Phi(\mathbf{r}) \\ X(\mathbf{r}) \end{pmatrix}. \tag{F.41}$$

F.2.2 Angular Part of the 2-Spinors

The two 2-spinors Φ and X can be expressed as linear combinations of eigenfunctions of simultaneously $\mathbf{j}^2$, j_z, $\mathbf{l}^2$ and $\mathbf{s}^2$ which is more easily seen by converting the term $\boldsymbol{\sigma}\mathbf{p}$ in a suitable form. We have[12]

$$\boldsymbol{\sigma}\cdot\mathbf{p} = \left(\frac{1}{r^2}\,\boldsymbol{\sigma}\cdot\mathbf{r}\;\;\boldsymbol{\sigma}\cdot\mathbf{r}\right)\boldsymbol{\sigma}\cdot\mathbf{p} = \frac{1}{r^2}\boldsymbol{\sigma}\cdot\mathbf{r}\;\;(\boldsymbol{\sigma}\cdot\mathbf{r}\;\;\boldsymbol{\sigma}\cdot\mathbf{p}) = \frac{1}{r^2}\boldsymbol{\sigma}\cdot\mathbf{r}\;\;(\mathbf{r}\cdot\mathbf{p} + i\,\boldsymbol{\sigma}\cdot\mathbf{l}) \tag{F.42}$$

or with $\hat{\mathbf{r}} = \frac{\mathbf{r}}{r}$

$$\boldsymbol{\sigma}\cdot\mathbf{p} = \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\left(\hat{\mathbf{r}}\cdot\mathbf{p} + \frac{i}{r}\,\boldsymbol{\sigma}\cdot\mathbf{l}\right). \tag{F.43}$$

Due to $\mathbf{j}^2 = (\mathbf{l}+\mathbf{s})^2 = \mathbf{l}^2 + 2\mathbf{s}\cdot\mathbf{l} + \mathbf{s}^2$, we have[13]

---

[12]We make use of $\boldsymbol{\sigma}\cdot\mathbf{a}\;\;\boldsymbol{\sigma}\cdot\boldsymbol{b} = \mathbf{a}\cdot\boldsymbol{b} + i\;\boldsymbol{\sigma}\cdot[\mathbf{a}\times\mathbf{b}]$ which means, inter alia, $\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\;\;\boldsymbol{\sigma}\cdot\hat{\mathbf{r}} = 1$. As usually, $\mathbf{r}\times\mathbf{p} = \mathbf{l}$.

[13]Remember $\mathbf{s} = \frac{\hbar}{2}\boldsymbol{\sigma}$.

$$2\mathbf{s}\cdot\mathbf{l} = \mathbf{j}^2 - \mathbf{l}^2 - \mathbf{s}^2. \tag{F.44}$$

The corresponding two-dimensional eigenspinors $\varphi_{jm_j}^{(+)}$ and $\varphi_{jm_j}^{(-)}$ (i.e., eigenfunctions of $\mathbf{j}^2$, j_z, $\mathbf{l}^2$ and $\mathbf{s}^2$) were derived in Sect. 16.5; they read in position representation (here, we note the angular part only)

$$\varphi_{jm_j}^{(+)} = \begin{pmatrix} \sqrt{\frac{l+m_j+1/2}{2l+1}}Y_l^{m_j-1/2} \\ \sqrt{\frac{l-m_j+1/2}{2l+1}}Y_l^{m_j+1/2} \end{pmatrix} \text{ for } j = l+\frac{1}{2}\ ;\ \varphi_{jm_j}^{(-)} = \begin{pmatrix} \sqrt{\frac{l-m_j+1/2}{2l+1}}Y_l^{m_j-1/2} \\ -\sqrt{\frac{l+m_j+1/2}{2l+1}}Y_l^{m_j+1/2} \end{pmatrix} \text{ for } j = l-\frac{1}{2}. \tag{F.45}$$

In comparison with Sect. 16.5, we have changed the sign for $\varphi_{jm_j}^{(-)}$ in order to arrive at the simple formulation (F.48).

The result may be written compactly as

$$\varphi_{jm_j}^{(\pm)} = \begin{pmatrix} \sqrt{\frac{l\pm m_j+1/2}{2l+1}}Y_l^{m_j-1/2} \\ \pm\sqrt{\frac{l\mp m_j+1/2}{2l+1}}Y_l^{m_j+1/2} \end{pmatrix} \text{ for } j = l \pm \frac{1}{2} \tag{F.46}$$

where the Y_m^l are spherical functions. We have $l = 0, 1, 2, \ldots,$ whereby $\varphi_{jm_j}^{(-)}$ exists only for $l > 0$. Note that for a given j, the two spinors have opposite parity,[14] since their l-values differ by one.

The eigenequations are

$$\begin{aligned} \mathbf{j}^2\varphi_{jm_j}^{(\pm)} &= \hbar^2 j\,(j+1)\,\varphi_{jm_j}^{(\pm)} \\ \mathbf{l}^2\varphi_{jm_j}^{(\pm)} &= \hbar^2 l\,(l+1)\,\varphi_{jm_j}^{(\pm)} \\ j_z\varphi_{jm_j}^{(\pm)} &= \hbar m_j\varphi_{jm_j}^{(\pm)} \\ \mathbf{s}^2\varphi_{jm_j}^{(\pm)} &= \tfrac{3}{4}\hbar^2\varphi_{jm_j}^{(\pm)}. \end{aligned} \tag{F.47}$$

We note in passing that

$$\varphi_{jm_j}^{(+)} = \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\varphi_{jm_j}^{(-)} \tag{F.48}$$

which again shows that $\varphi_{jm_j}^{(+)}$ and $\varphi_{jm_j}^{(-)}$ have opposite parity.[15]

F.2.3 From 2-Spinors to 4-Spinor

Thus, the general expression for the four-spinor Ψ for given values of j and m_j reads[16]

---

[14]Remind that the parity of Y_l^m is given by $(-1)^l$.

[15]Note that in spherical coordinates, $\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}$ only contains the angle variables ϑ and φ.

[16]The factors i and $\frac{1}{r}$ are chosen in order to simplity the following calculations.

$$\Psi_{jm_j} = \frac{1}{r}\begin{pmatrix} iG_j^{(+)}(r)\,\varphi_{jm_j}^{(+)} + iG_j^{(-)}(r)\,\varphi_{jm_j}^{(-)} \\ F_j^{(+)}(r)\,\varphi_{jm_j}^{(-)} + F_j^{(-)}(r)\,\varphi_{jm_j}^{(+)} \end{pmatrix}. \tag{F.49}$$

We can split this expression into two solutions with defined parity. Due to the (F.40) and (F.48), they read

$$\Psi_{jm_j}^{(+)} = \frac{1}{r}\begin{pmatrix} iG_j^{(+)}(r)\,\varphi_{jm_j}^{(+)} \\ F_j^{(+)}(r)\,\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\varphi_{jm_j}^{(+)} \end{pmatrix}\ ;\ \Psi_{jm_j}^{(-)} = \frac{1}{r}\begin{pmatrix} iG_j^{(-)}(r)\,\varphi_{jm_j}^{(-)} \\ F_j^{(-)}(r)\,\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\varphi_{jm_j}^{(-)} \end{pmatrix} \tag{F.50}$$

which usually is written compactly as

$$\Psi_{jm_j}^{l} = \frac{1}{r}\begin{pmatrix} iG_{jl}(r)\,\varphi_{jm_j}^{l} \\ F_{jl}(r)\,\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\varphi_{jm_j}^{l} \end{pmatrix}\ ;\ j = l \pm \frac{1}{2}. \tag{F.51}$$

Note that $\Psi_{jm_j}^{l}$ has parity $(-1)^l$.

Inserting this expression into (F.38) results in

$$\begin{array}{c} E\frac{G_{jl}}{r}\varphi_{jm_j}^{l} = -ic\,\boldsymbol{\sigma}\cdot\mathbf{p}\,\frac{F_{jl}}{r}\,\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\varphi_{jm_j}^{l} + V(r)\,\frac{G_{jl}}{r}\varphi_{jm_j}^{l} + m_0c^2\frac{G_{jl}}{r}\varphi_{jm_j}^{l} \\ E\frac{F_{jl}}{r}\,\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\varphi_{jm_j}^{l} = ic\,\boldsymbol{\sigma}\cdot\mathbf{p}\,\frac{G_{jl}}{r}\varphi_{jm_j}^{l} + V(r)\,\frac{F_{jl}}{r}\,\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\varphi_{jm_j}^{l} - m_0c^2\frac{F_{jl}}{r}\,\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\varphi_{jm_j}^{l} \end{array}. \tag{F.52}$$

We now consider the action of the operators $\boldsymbol{\sigma}\cdot\mathbf{p}$ and $\boldsymbol{\sigma}\cdot\mathbf{p}\ \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}$ on the eigenfunctions. With (F.43), we have for a general function $H(\mathbf{r})$

$$\boldsymbol{\sigma}\cdot\mathbf{p}\,H(\mathbf{r}) = \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\left(\hat{\mathbf{r}}\cdot\mathbf{p} + \frac{i}{r}\,\boldsymbol{\sigma}\cdot\mathbf{l}\right)H(\mathbf{r}) = \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\left(\hat{\mathbf{r}}\cdot\mathbf{p} + \frac{2i}{\hbar r}\mathbf{s}\cdot\mathbf{l}\right)H(\mathbf{r}). \tag{F.53}$$

Using $\hat{\mathbf{r}}\cdot\mathbf{p}{=}\frac{\hbar}{i}\frac{\partial}{\partial r}$ and $2\mathbf{s}\cdot\mathbf{l} = \mathbf{j}^2 - \mathbf{l}^2 - \mathbf{s}^2$ yields

$$\begin{array}{c} \boldsymbol{\sigma}\cdot\mathbf{p}\,H(r)\,\varphi_{jm_j}^{l} = \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\left(\frac{\hbar}{i}\frac{\partial}{\partial r} + \frac{i}{\hbar r}\left[\mathbf{j}^2 - \mathbf{l}^2 - \mathbf{s}^2\right]\right)H(r)\,\varphi_{jm_j}^{l} = \\ = \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\left(\frac{\hbar}{i}\frac{\partial}{\partial r} + \frac{i\hbar}{r}\left[j(j+1) - l(l+1) - \frac{3}{4}\right]\right)H(r)\,\varphi_{jm_j}^{l}. \end{array} \tag{F.54}$$

With

$$j(j+1) - l(l+1) - \frac{3}{4} = \begin{cases} j - \frac{1}{2} = -1 + \left(j + \frac{1}{2}\right) \\ -j - \frac{3}{2} = -1 - \left(j + \frac{1}{2}\right) \end{cases} \text{for } j = l \pm \frac{1}{2} \tag{F.55}$$

we arrive at

$$\begin{array}{c} \boldsymbol{\sigma}\cdot\mathbf{p}\,H(r)\,\varphi_{jm_j}^{l} = \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\left(\frac{\hbar}{i}\frac{\partial}{\partial r} + \frac{i\hbar}{r}\left[j(j+1) - l(l+1) - \frac{3}{4}\right]\right)H(r)\,\varphi_{jm_j}^{l} = \\ = \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\frac{\hbar}{ir}\left[r\frac{\partial H(r)}{\partial r} + \left(1 \mp \left(j + \frac{1}{2}\right)\right)H(r)\right]\varphi_{jm_j}^{l}\ \text{for } j = l \pm \frac{1}{2}. \end{array} \tag{F.56}$$

In addition, we need $\boldsymbol{\sigma}\cdot\mathbf{p}\,H(r)\,\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\,\varphi_{jm_j}^{l}$:

$$\begin{aligned} &\boldsymbol{\sigma}\cdot\mathbf{p}\, H(r)\, \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\, \varphi^{l}_{jm_j} = \boldsymbol{\sigma}\cdot\mathbf{p}\ \ \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\, H(r)\varphi^{l}_{jm_j} = \\ &= \left[\mathbf{p}\cdot\hat{\mathbf{r}} + i\boldsymbol{\sigma}\left(\mathbf{p}\times\hat{\mathbf{r}}\right)\right] H(r)\varphi^{l}_{jm_j} = \left(-\tfrac{2i\hbar}{r} + \hat{\mathbf{r}}\cdot\mathbf{p} - \tfrac{i}{r}\,\boldsymbol{\sigma}\cdot\mathbf{L}\right) H(r)\varphi^{l}_{jm_j}. \end{aligned} \tag{F.57}$$

Evaluating $\hat{\mathbf{r}}\cdot\mathbf{p}$ and $\frac{i}{r}\,\boldsymbol{\sigma}\cdot\mathbf{L}$ yields

$$\begin{aligned} \boldsymbol{\sigma}\cdot\mathbf{p}\, H(r)\, \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\, \varphi^{l}_{jm_j} &= \left(-\tfrac{2i\hbar}{r} + \tfrac{\hbar}{i}\tfrac{\partial}{\partial r} - \tfrac{i\hbar}{r}\left[j\,(j+1) - l\,(l+1) - \tfrac{3}{4}\right]\right) H(r)\varphi^{l}_{jm_j} = \\ &= \left(-\tfrac{2i\hbar}{r} + \tfrac{\hbar}{i}\tfrac{\partial}{\partial r} + \tfrac{i\hbar}{r}\left[1 \mp \left(j+\tfrac{1}{2}\right)\right]\right) H(r)\varphi^{l}_{jm_j} \end{aligned} \tag{F.58}$$

or

$$\boldsymbol{\sigma}\cdot\mathbf{p}\, H(r)\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\, \varphi^{l}_{jm_j} = \frac{\hbar}{ir}\left[r\frac{\partial}{\partial r} + 1 \pm \left(j+\frac{1}{2}\right)\right] H(r)\varphi^{l}_{jm_j} \ \text{ for } j = l \pm \frac{1}{2}. \tag{F.59}$$

Inserting the results in (F.52) yields finally

$$\begin{aligned} -\tfrac{1}{c\hbar}\left[E + \tfrac{1}{4\pi\varepsilon_0}\tfrac{Ze^2}{r} - m_0c^2\right] G_{jl} &= \tfrac{\partial}{\partial r}F_{jl} \pm \tfrac{j+\frac{1}{2}}{r}F_{jl} \\ \tfrac{1}{c\hbar}\left[E + \tfrac{1}{4\pi\varepsilon_0}\tfrac{Ze^2}{r} + m_0c^2\right] F_{jl} &= \tfrac{\partial}{\partial r}G_{jl} \mp \tfrac{j+\frac{1}{2}}{r}G_{jl}. \end{aligned} \tag{F.60}$$

F.2.4 Coupled Radial Equations, Solution

To get rid of the many constants, we insert the fine-structure constant[17] α

$$\alpha = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{c\hbar} \tag{F.61}$$

and introduce the abbreviations

$$k = \pm\left(j+\frac{1}{2}\right)\ ;\ \alpha_1 = \frac{m_0c^2 + E}{c\hbar}\ ;\ \alpha_2 = \frac{m_0c^2 - E}{c\hbar}\ ;\ \tau = \sqrt{\alpha_1\alpha_2}\ ;\ \rho = \tau r\ ;\ \gamma = Z\alpha. \tag{F.62}$$

Thus, (F.60) are written as

$$\begin{aligned} \left(\tfrac{\mathrm{d}}{\mathrm{d}\rho} + \tfrac{k}{\rho}\right) F_{jl} - \left(\tfrac{\alpha_2}{\tau} - \tfrac{\gamma}{\rho}\right) G_{jl} &= 0 \\ \left(\tfrac{\mathrm{d}}{\mathrm{d}\rho} - \tfrac{k}{\rho}\right) G_{jl} - \left(\tfrac{\alpha_1}{\tau} + \tfrac{\gamma}{\rho}\right) F_{jl} &= 0. \end{aligned} \tag{F.63}$$

From now on, the method is essentially the same as in the case of the Hydrogen atom of the Schrödinger equation. We just sketch the essential steps of the approach,

[17] Do not confuse the fine structure constant α and the Dirac matrices $\boldsymbol{\alpha}$.

leaving the detailed calculation to the reader (it is just too comprehensive to be performed in an appendix).

First, one can show that the physically meaningful solutions of (F.63) have an asymptotical behavior proportional to $e^{-\rho}$. In view of this, one separates this behavior by choosing

$$F_{jl}(\rho) = e^{-\rho} f(\rho) \;\; ; \; G_{jl}(\rho) = e^{-\rho} g(\rho) \, . \tag{F.64}$$

The solutions of the resulting differential equations for f and g are searched in form of ρ^s times power series. Demanding regularity of the functions at $\rho = 0$ leads to $s = \sqrt{k^2 - \gamma^2}$. Furthermore, comparison of the coefficients of equal powers in the power series yields recursion relations between these coefficients. The determination of the coefficients leads to power series with an asymptotic behavior $\sim e^{2\rho}$ which would according to (F.64) yield non-normalizable functions. This unphysical behavior can only be avoided, if the power series stop at a certain power N, i.e., if they are polynomials.[18] The termination condition reads

$$E = m_0 c^2 \left[1 + \frac{\gamma^2}{(s+N)^2} \right]^{-\frac{1}{2}} . \tag{F.65}$$

With the *main quantum number n*

$$n = N + j + \frac{1}{2} \tag{F.66}$$

the energy levels for Coulomb interaction are given by:

$$E_{nj} = m_0 c^2 \left[1 + \left(\frac{Z\alpha}{n - \left(j + \frac{1}{2}\right) + \sqrt{\left(j + \frac{1}{2}\right)^2 - (Z\alpha)^2}} \right)^2 \right]^{-\frac{1}{2}} \tag{F.67}$$

with

$$n = 1, 2, \ldots, \infty \;\; ; \; 0 < j + \frac{1}{2} \leq n \;\; ; \; 0 \leq l \leq n - 1 \; , \; j = l \pm \frac{1}{2} . \tag{F.68}$$

For a discussion of this spectrum see Sect. 19.3, Vol. 2.

F.3 Exercises and Solutions

1. Calculate $\mathbf{r} \cdot \mathbf{p}$, $\mathbf{p} \cdot \mathbf{r}$ and $\boldsymbol{\sigma} \cdot \mathbf{p}$
2. Prove (F.48), i.e.

[18] N is called radial quantum number.

$$\varphi_{jm_j}^{(+)} = \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\, \varphi_{jm_j}^{(-)}. \tag{F.69}$$

Solution: We have

$$\hat{\mathbf{r}} = \frac{1}{r}(x, y, z) = \left(-\frac{1}{2}\sqrt{\frac{8\pi}{3}}\left(Y_1^1 - Y_1^{-1}\right), -\frac{1}{2i}\sqrt{\frac{8\pi}{3}}\left(Y_1^1 + Y_1^{-1}\right), \sqrt{\frac{4\pi}{3}}Y_1^0\right). \tag{F.70}$$

It follows

$$\begin{aligned}\boldsymbol{\sigma}\cdot\hat{\mathbf{r}} &= \frac{1}{r}\left(\sigma_x x + \sigma_y y + \sigma_z z\right) = \frac{1}{r}\left[\begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix} + \begin{pmatrix} 0 & -iy \\ iy & 0 \end{pmatrix} + \begin{pmatrix} z & 0 \\ 0 & -z \end{pmatrix}\right] \\ &= \frac{1}{r}\begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix}\end{aligned} \tag{F.71}$$

or

$$\boldsymbol{\sigma}\cdot\hat{\mathbf{r}} = \begin{pmatrix} \sqrt{\frac{4\pi}{3}}Y_1^0 & \sqrt{\frac{8\pi}{3}}Y_1^{-1} \\ -\sqrt{\frac{8\pi}{3}}Y_1^1 & -\sqrt{\frac{4\pi}{3}}Y_1^0 \end{pmatrix} = \sqrt{\frac{4\pi}{3}}\begin{pmatrix} Y_1^0 & \sqrt{2}Y_1^{-1} \\ -\sqrt{2}Y_1^1 & -Y_1^0 \end{pmatrix}. \tag{F.72}$$

Appendix G
The Lenz Vector

In this section, we want to derive the spectrum of the hydrogen atom in an algebraic manner; the analytic derivation can be found in Chap. 17 and Appendix F, Vol. 2. Here, we use the fact that the *Lenz vector*[19] is a constant of the motion (and thus a conserved quantity); this additional constant is, moreover, responsible for the high degree of degeneracy in the energy spectrum of the hydrogen atom.

The eigenfunctions of the hydrogen atom can be represented algebraically as well; this is given in some textbooks on quantum mechanics (see e.g. Schwabl, Annex C., p. 400).

G.1 In Classical Mechanics

In classical mechanics, the Lenz vector is defined as:

$$\mathbf{\Lambda}_{ClM} = \frac{1}{m\gamma} (\mathbf{L} \times \mathbf{p}) + \frac{\mathbf{r}}{r}. \tag{G.1}$$

For the motion of a particle in a Coulomb field (or a Kepler field) $V(r) = -\gamma/r$, it is, in addition to the energy and the angular momentum, a further conserved quantity. Its magnitude is equal to the eccentricity of the elliptical orbit. Its conservation means that this orbital ellipse is not rotating, so there is no perihelion motion. For other potentials, this is generally[20] not the case; there, one finds *rosette orbits*.

[19] Also called the Laplace–Runge–Lenz vector, Runge–Lenz vector, etc. and, especially in quantum mechanics, the Runge–Lenz–Pauli operator.

[20] Aside from the Coulomb potential, only the potential of the harmonic oscillator, $V \sim r^2$, leads to closed elliptical orbits.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

G.2 In Quantum Mechanics

For the translation into quantum mechanics, we have to symmetrize:

$$\mathbf{\Lambda} = \frac{1}{2m\gamma}\left(\mathbf{L}\times\mathbf{p}-\mathbf{p}\times\mathbf{L}\right) + \frac{\mathbf{r}}{r}. \tag{G.2}$$

The Hamiltonian reads

$$H = \frac{p^2}{2m} - \frac{\gamma}{r}. \tag{G.3}$$

$\mathbf{\Lambda}$ is a Hermitian vector operator that commutes with H (i.e. it represents a vectorial conserved quantity) and is orthogonal[21] to $\mathbf{L}$:

$$\mathbf{\Lambda} = \mathbf{\Lambda}^\dagger;\ \ [\mathbf{\Lambda}, H] = 0;\ \ \mathbf{\Lambda}\cdot\mathbf{L} = \mathbf{L}\cdot\mathbf{\Lambda} = 0. \tag{G.4}$$

In principle, it is technically rather easy to prove these and the other statements that follow, but often it is also a lengthy procedure. Therefore, the proofs are left to the exercises.

For $\mathbf{\Lambda}^2$, we obtain[22]:

$$\mathbf{\Lambda}^2 = \frac{2H}{m\gamma^2}\left(\mathbf{L}^2 + \hbar^2\right) + 1. \tag{G.5}$$

We restrict ourselves to negative energies $-|E|$ (and thus to bound states; in principle, the reasoning could be extended to scattering states). With the rescaling[23]

$$\mathbf{R} = \sqrt{\frac{m\gamma^2}{2|E|}}\mathbf{\Lambda}, \tag{G.6}$$

it follows that

$$\mathbf{R}^2 + \mathbf{L}^2 + \hbar^2 = -\frac{m\gamma^2}{2E}. \tag{G.7}$$

Finally, we introduce two generalized angular-momentum operators:

$$\mathbf{J}_1 = \frac{1}{2}\left(\mathbf{L} + \mathbf{R}\right);\ \mathbf{J}_2 = \frac{1}{2}\left(\mathbf{L} - \mathbf{R}\right). \tag{G.8}$$

[21] $\mathbf{\Lambda}$ is a polar vector, $\mathbf{L}$ an axial vector.

[22] We recall that for a central potential, $\left[H, \mathbf{L}^2\right] = 0$.

[23] The vector operator $\mathbf{R}$ introduced here is not to be confused with the center-of-mass vector.

They satisfy the equations

$$[\mathbf{J}_1, \mathbf{J}_2] = 0;\ \mathbf{J}_1^2 = \mathbf{J}_2^2. \tag{G.9}$$

With (G.8) , we can write (G.7) in the form:

$$2\left(\mathbf{J}_1^2 + \mathbf{J}_2^2\right) + \hbar^2 = -\frac{m\gamma^2}{2E}, \tag{G.10}$$

and because of (G.9), it follows that

$$4\mathbf{J}_1^2 + \hbar^2 = -\frac{m\gamma^2}{2E}. \tag{G.11}$$

Because $\mathbf{J}_1$ is a generalized angular-momentum operator, its eigenvalues have the form $\hbar^2 j\,(j + 1)$, where j is a positive integer or half-integer. (This operator has a different symmetry from the angular-momentum operators considered previously; therefore its eigenvalues can take on both integer and half-integer values.)

Hence, it follows that

$$4\hbar^2 j\,(j + 1) + \hbar^2 = \frac{m\gamma^2}{2\,|E|} \tag{G.12}$$

or

$$E = -\frac{m\gamma^2}{2\hbar^2\,(2j + 1)^2};\ j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots \tag{G.13}$$

If we identify the numbers $2j + 1$ with the principal quantum number n, we obtain the the familiar form of the energy levels of the hydrogen atom.

G.3 General Theorems on Vector Operators

For the manipulations in the exercises, we compile here a few facts about commutators and vector operators:

G.3.1 General Commutator Relations

We need, among others, the general commutator relations

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \tag{G.14}$$

(Jacobi identity) and

$$[A, BC] = [A, B]\, C \;+\; B\, [A, C]$$
$$[AB, C] = A\, [B, C] \;+\; [A, C]\, B. \tag{G.15}$$

G.3.2 Vector Operators

An operator is a vector operator iff

$$\left[L_j, A_k\right] = i\hbar \sum_m \varepsilon_{jkm} A_m; \;\; \left[\mathbf{L}^2, \mathbf{A}^2\right] = 0. \tag{G.16}$$

Here, **L** is the angular-momentum operator and ε_{jkm} is the Levi-Civita tensor (Levi-Civita symbol, permutation tensor, antisymmetric symbol, alternating symbol; see Appendix F, Vol. 1).[24]

For two vector operators **B** and **C**, it holds that:

$$[\mathbf{L}, \mathbf{B} \cdot \mathbf{C}] = 0. \tag{G.17}$$

This is due to

$$[L_i, \mathbf{B} \cdot \mathbf{C}] = \sum\nolimits_j \left[L_i, B_j C_j\right] = \sum\nolimits_j \left[L_i, B_j\right] C_j + B_j \left[L_i, C_j\right]$$
$$= i\hbar \sum\nolimits_j \sum\nolimits_m \varepsilon_{ijm} \left(B_m C_j + B_j C_m\right) = 0. \tag{G.18}$$

The last step follows because of $\varepsilon_{1jm} = -\varepsilon_{1mj}$:

$$\sum_{jm} \varepsilon_{ijm} \left(B_m C_j + B_j C_m\right) = \sum_{jm} B_m C_j \left(\varepsilon_{ijm} \;+\; \varepsilon_{imj}\right) = 0. \tag{G.19}$$

In addition, for a vector operator, we have:

$$\mathbf{A} \times \mathbf{L} = -\mathbf{L} \times \mathbf{A} \;+\; 2i\hbar\mathbf{A}; \;\; \mathbf{L} \times \mathbf{A} = -\mathbf{A} \times \mathbf{L} \;+\; 2i\hbar\mathbf{A}. \tag{G.20}$$

[24]Typically, in this context, the Einstein summation convention is used, according to which one sums over repeated indices without noting this explicitly. Since we rarely use the Leci-Civita tensor, we write out the summation sign (cf. Appendix F, Vol. 1). Instead of using the usual notation, for example

$$\varepsilon_{ijk}\varepsilon_{mnk} = \delta_{im}\delta_{jn} - \delta_{\text{in}}\delta_{jm},$$

we write

$$\sum_k \varepsilon_{ijk}\varepsilon_{mnk} = \delta_{im}\delta_{jn} - \delta_{\text{in}}\delta_{jm}.$$

Proof:

$$\begin{aligned}(\mathbf{A}\times\mathbf{L})_i &= \sum\nolimits_{jk}\varepsilon_{ijk}A_jL_k = \sum\nolimits_{jk}\varepsilon_{ijk}\left(L_kA_j - \hbar\sum\nolimits_m \varepsilon_{jkm}A_m\right)\\ &= \sum\nolimits_{jk}\varepsilon_{ijk}L_kA_j - \sum\nolimits_{jk}\varepsilon_{ijk}\sum\nolimits_m \varepsilon_{jkm}A_m = -\left(\mathbf{L}\times\mathbf{A}\right) + 2i\hbar\mathbf{A},\end{aligned} \tag{G.21}$$

due to

$$\sum_{jk}\varepsilon_{ijk}\varepsilon_{jkm} = -2\delta_{mi}. \tag{G.22}$$

Furthermore, vector operators satisfy generally the equation

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = (\mathbf{A}\times\mathbf{B})\cdot\mathbf{C}. \tag{G.23}$$

In particular, for the momentum and the position, we have

$$\mathbf{r}\cdot\mathbf{L} = \mathbf{L}\cdot\mathbf{r} = 0;\ \ \mathbf{p}\cdot\mathbf{L} = \mathbf{L}\cdot\mathbf{p} = 0. \tag{G.24}$$

G.4 Exercises

1. Show that $\mathbf{\Lambda}$ is a Hermitian vector operator which commutes with $H = \frac{\mathbf{p}^2}{2m} - \frac{\gamma}{r}$ and satisfies the equation $\mathbf{\Lambda}\cdot\mathbf{L} = \mathbf{L}\cdot\mathbf{\Lambda} = 0$.
   Solution:

   (a) For the Hermiticity, we need only consider the term $(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L})$ which, due to the symmetrization (and because $\mathbf{L}$ and $\mathbf{p}$ are Hermitian), is automatically Hermitian. We see this explicitly for e.g. the x or 1 component:

$$\begin{aligned}(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L})_1 &= L_2p_3 - L_3p_2 - p_2L_3 + p_3L_2\\ (\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L})_1^\dagger &= p_3L_2 - p_2L_3 - L_3p_2 + L_2p_3\\ &= (\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L})_1\,.\end{aligned} \tag{G.25}$$

   (b) Concerning the question of the vector operator, we realize first that

$$\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L} = \frac{i}{\hbar}\left[\mathbf{L}^2, \mathbf{p}\right]. \tag{G.26}$$

   We calculate this following the example of the x or 1 component:

$$\left[\mathbf{L}^2, p_1\right] = \sum_i\left[L_i^2, p_1\right] = \sum_i L_i\left[L_i, p_1\right] + \left[L_i, p_1\right]L_i \tag{G.27}$$

where we have used (G.15). Since the momentum is a vector operator, we can use (G.16) and obtain explicitly

$$\begin{aligned}\left[\mathbf{L}^2, p_1\right] &= \sum_i \sum_m L_i i\hbar\varepsilon_{i1m} p_m + i\hbar\varepsilon_{i1m} p_m L_i \\ &= L_2 i\hbar\varepsilon_{213} p_3 + i\hbar\varepsilon_{213} p_3 L_2 + L_3 i\hbar\varepsilon_{312} p_2 + i\hbar\varepsilon_{312} p_2 L_3.\end{aligned} \tag{G.28}$$

With the corresponding values for the Levi-Civita tensor, it follows that

$$\begin{aligned}\left[\mathbf{L}^2, p_1\right] &= i\hbar\left(-L_2 p_3 - p_3 L_2 + L_3 p_2 + p_2 L_3\right) \\ &= -i\hbar\left(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}\right)_1\end{aligned} \tag{G.29}$$

and thus (G.26). Hence, we can write the Lenz vector $\mathbf{\Lambda} = \frac{1}{2m\gamma}\left(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}\right) + \hat{\mathbf{r}}$ also as

$$\mathbf{\Lambda} = \frac{i}{2m\gamma\hbar}\left[\mathbf{L}^2, \mathbf{p}\right] + \hat{\mathbf{r}}. \tag{G.30}$$

Now we can treat the question of the vector operator. We have to show that:

$$\left[L_j, \Lambda_k\right] = i\hbar\sum_m \varepsilon_{jkm}\Lambda_m. \tag{G.31}$$

Since $\hat{\mathbf{r}}$ is a vector operator, we can confine ourselves to $\left[\mathbf{L}^2, \mathbf{p}\right]$, i.e. we have to show that

$$\left[L_j, \left[\mathbf{L}^2, p_k\right]\right] = i\hbar\sum_m \varepsilon_{jkm}\left[\mathbf{L}^2, p_m\right]. \tag{G.32}$$

With the Jacobi identity (G.14), we obtain

$$\left[\mathbf{L}^2, \left[L_j, p_k\right]\right] = i\hbar\sum_m \varepsilon_{jkm}\left[\mathbf{L}^2, p_m\right]. \tag{G.33}$$

Since the momentum is a vector operator, we find with $\left[L_j, p_k\right] = i\hbar\sum_m \varepsilon_{jkm} p_m$ from the last equation that:

$$\left[\mathbf{L}^2, i\hbar\sum_m \varepsilon_{jkm} p_m\right] = i\hbar\sum_m \varepsilon_{jkm}\left[\mathbf{L}^2, p_m\right], \tag{G.34}$$

with which the vector operator character of $\mathbf{\Lambda}$ is proven.

(c) The question of the commutator is next: We have to show that

$$[H, \mathbf{\Lambda}] = \left[H, \frac{i}{2m\gamma\hbar}\left[\mathbf{L}^2, \mathbf{p}\right] + \hat{\mathbf{r}}\right] = 0. \tag{G.35}$$

We will do this again step by step and in detail. First, we have

$$\begin{aligned}[H, \mathbf{\Lambda}] &= \frac{i}{2m\gamma\hbar}\left[H, \left[\mathbf{L}^2, \mathbf{p}\right]\right] + \left[H, \hat{\mathbf{r}}\right] \\ &= -\frac{i}{2m\gamma\hbar}\left[\mathbf{L}^2, \left[\mathbf{p}, H\right]\right] + \left[H, \hat{\mathbf{r}}\right]\end{aligned} \tag{G.36}$$

where we have used the fact that for a central potential, $\left[H, \mathbf{L}^2\right] = 0$. We insert $H = \frac{p^2}{2m} - \frac{\gamma}{r}$ and obtain

$$[H, \mathbf{\Lambda}] = \frac{i}{2m\hbar}\left[\mathbf{L}^2, \left[\mathbf{p}, \frac{1}{r}\right]\right] + \frac{1}{2m}\left[\mathbf{p}^2, \hat{\mathbf{r}}\right]. \tag{G.37}$$

Calculating the commutator $\left[\mathbf{p}, \frac{1}{r}\right]$ yields

$$\left[\mathbf{p}, \frac{1}{r}\right] = -\frac{\hbar}{i}\frac{\hat{\mathbf{r}}}{r^2}. \tag{G.38}$$

It follows that

$$[H, \mathbf{\Lambda}] = -\frac{1}{2m}\left[\mathbf{L}^2, \frac{\hat{\mathbf{r}}}{r^2}\right] + \frac{1}{2m}\left[\mathbf{p}^2, \hat{\mathbf{r}}\right]. \tag{G.39}$$

We know that $\mathbf{L}^2$ contains derivatives only with respect to the angles; therefore we can write

$$[H, \mathbf{\Lambda}] = \frac{1}{2m}\left[\mathbf{p}^2, \hat{\mathbf{r}}\right] - \frac{1}{2m}\left[\frac{\mathbf{L}^2}{r^2}, \hat{\mathbf{r}}\right]. \tag{G.40}$$

From the representation of the Laplacian in spherical coordinates, we obtain:

$$\mathbf{p}^2 = p_r^2 + \frac{\mathbf{L}^2}{r^2}. \tag{G.41}$$

It follows that

$$[H, \mathbf{\Lambda}] = \frac{1}{2m}\left[p_r^2 + \frac{\mathbf{L}^2}{r^2}, \hat{\mathbf{r}}\right] - \frac{1}{2m}\left[\frac{\mathbf{L}^2}{r^2}, \hat{\mathbf{r}}\right] = \frac{1}{2m}\left[p_r^2, \hat{\mathbf{r}}\right] = 0. \tag{G.42}$$

The last equals sign applies since p_r contains only derivatives with respect to r, while in $\hat{\mathbf{r}}$, only angles occur.

(d) Finally, we show that $\boldsymbol{\Lambda} \cdot \mathbf{L} = \mathbf{L} \cdot \boldsymbol{\Lambda} = 0$. We start with $\boldsymbol{\Lambda} = \frac{1}{2m\gamma} (\mathbf{L} \times \mathbf{p} - \mathbf{p} \times \mathbf{L}) + \frac{\mathbf{r}}{r}$ and find:

$$\boldsymbol{\Lambda} \cdot \boldsymbol{L} = \frac{1}{2m\gamma} (\mathbf{L} \times \mathbf{p} - \mathbf{p} \times \mathbf{L}) \cdot \mathbf{L} + \hat{\mathbf{r}} \cdot \mathbf{L}. \tag{G.43}$$

From Chap. 16, we know that $\hat{\mathbf{r}} \cdot \mathbf{L} = \mathbf{L} \cdot \hat{\mathbf{r}} = 0$. (Hint: Write the operator equation out in coordinates and examine it in detail). So we have to show that

$$(\mathbf{L} \times \mathbf{p} - \mathbf{p} \times \mathbf{L}) \cdot \mathbf{L} = \mathbf{L} \cdot (\mathbf{L} \times \mathbf{p} - \mathbf{p} \times \mathbf{L}) = 0. \tag{G.44}$$

We consider here only $\mathbf{L} \cdot (\mathbf{L} \times \mathbf{p} - \mathbf{p} \times \mathbf{L})$; the treatment of the other term proceeds analogously. First, we rewrite using (G.20) and obtain

$$\mathbf{L} \cdot (\mathbf{L} \times \mathbf{p} - \mathbf{p} \times \mathbf{L}) = \mathbf{L} \cdot (2\mathbf{L} \times \mathbf{p} - \mathbf{2}i\hbar\mathbf{p}.) \tag{G.45}$$

We can rearrange this using (G.23):

$$\mathbf{L} \cdot (2\mathbf{L} \times \mathbf{p} - \mathbf{2}i\hbar\mathbf{p}) = 2\,(\mathbf{L} \times \mathbf{L}) \cdot \mathbf{p} - \mathbf{2}i\hbar\mathbf{L} \cdot \mathbf{p} = 0 \tag{G.46}$$

because of $\mathbf{L} \cdot \mathbf{p} = 0$.

For practice, we perform the calculation once more using the Levi-Civita tensor:

$$(\mathbf{L} \times \mathbf{p} - \mathbf{p} \times \mathbf{L}) \cdot \mathbf{L} = \sum_{ijk} \varepsilon_{ijk} \left(L_j p_k - p_j L_k \right) L_i. \tag{G.47}$$

Here, we insert

$$\left[L_j, p_k \right] = i\hbar \sum_m \varepsilon_{jkm} p_m \tag{G.48}$$

and obtain

$$\begin{aligned}(\mathbf{L} \times \mathbf{p} - \mathbf{p} \times \mathbf{L}) \cdot \mathbf{L} &= \sum_{ijk} \varepsilon_{ijk} \left[i\hbar \sum_m \varepsilon_{jkm} p_m - p_k L_j - p_j L_k \right] L_i \\ &= i\hbar \sum_{ijkm} \varepsilon_{ijk} \varepsilon_{jkm} p_m L_i - \sum_{ijk} \varepsilon_{ijk} \left(p_k L_j + p_j L_k \right) L_i.\end{aligned} \tag{G.49}$$

We rewrite the last summand:

$$\sum_{ijk}\varepsilon_{ijk}\left(p_kL_j + p_jL_k\right)L_i = \sum_{ijk}\left(\varepsilon_{ijk} + \varepsilon_{ikj}\right)p_kL_jL_kL_i = 0$$
$$\text{due to } \varepsilon_{ijk} = -\varepsilon_{ikj}. \tag{G.50}$$

For the first summand, we have using (G.22):

$$\sum_{ijkm}\varepsilon_{ijk}\varepsilon_{jkm}p_mL_i = -2\sum_{im}\delta_{mi}p_mL_i$$
$$= -2\sum_i p_iL_i = -2\mathbf{p}\cdot\mathbf{L} = 0. \tag{G.51}$$

2. Prove the equation

$$\boldsymbol{\Lambda}^2 = \frac{2H}{m\gamma^2}\left(\mathbf{L}^2 + \hbar^2\right) + 1. \tag{G.52}$$

Solution: With the abbreviation

$$g = \frac{1}{2m\gamma}, \tag{G.53}$$

we can write the square of the Lenz vector

$$\boldsymbol{\Lambda} = \frac{1}{2m\gamma}\left(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}\right) + \frac{\mathbf{r}}{r} \tag{G.54}$$

as

$$\boldsymbol{\Lambda}^2 = g^2\left(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}\right)\left(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}\right)$$
$$+ g\left[\left(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}\right)\frac{\mathbf{r}}{r} + \frac{\mathbf{r}}{r}\left(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}\right)\right] + 1. \tag{G.55}$$

We treat the terms $\sim g^2$ and $\sim g$ separately.

(a) The terms $\sim g^2$ are

$$(\mathbf{L}\times\mathbf{p})(\mathbf{L}\times\mathbf{p}) - (\mathbf{p}\times\mathbf{L})(\mathbf{L}\times\mathbf{p}) - (\mathbf{L}\times\mathbf{p})(\mathbf{p}\times\mathbf{L}) + (\mathbf{p}\times\mathbf{L})(\mathbf{p}\times\mathbf{L})$$
$$= p^2L^2 - (-p^2L^2 - 4\hbar^2p^2) - (-p^2L^2) + p^2L^2 = 4p^2L^2 + 4\hbar^2p^2. \tag{G.56}$$

To realize this, we first consider the last term:

$$(\mathbf{p}\times\mathbf{L})(\mathbf{p}\times\mathbf{L}) = \sum_i\sum_{jk,mn}\varepsilon_{ijk}p_jL_k\varepsilon_{imn}p_mL_n$$
$$= \sum_{jk}p_jL_kp_jL_k - p_jL_kp_kL_j$$

$$
\begin{aligned}
&= \sum_{jk} p_j \left(p_j L_k - i\hbar \sum_m \varepsilon_{jkm} p_m \right) L_k - p_j \left(\mathbf{L}\cdot\mathbf{p}\right) L_j \\
&= \sum_{jk} \left(p_j^2 L_k^2 - i\hbar p_j \sum_m \varepsilon_{jkm} p_m L_k \right) \\
&= \mathbf{p}^2\mathbf{L}^2 - i\hbar \sum_k \left(\sum_{jm} \varepsilon_{jkm} p_j p_m \right) L_k \\
&= \mathbf{p}^2\mathbf{L}^2 + i\hbar \sum_k \left(\mathbf{p}\times\mathbf{p}\right)_k L_k = \mathbf{p}^2\mathbf{L}^2 = p^2 L^2, \qquad \text{(G.57)}
\end{aligned}
$$

where we have used $\mathbf{L}\cdot\mathbf{p} = 0$.

With this equation and with (G.20) and (G.23), the other terms can be calculated. We have

$$
\begin{aligned}
&(\mathbf{L}\times\mathbf{p})(\mathbf{p}\times\mathbf{L}) = (-\mathbf{p}\times\mathbf{L} + 2i\hbar\mathbf{p})(\mathbf{p}\times\mathbf{L}) \\
&\quad = -(\mathbf{p}\times\mathbf{L})(\mathbf{p}\times\mathbf{L}) + 2i\hbar\mathbf{p}(\mathbf{p}\times\mathbf{L}) = -p^2L^2 + 2i\hbar(\mathbf{p}\times\mathbf{p})\mathbf{L} = -p^2L^2 \qquad \text{(G.58)}
\end{aligned}
$$

and

$$
\begin{aligned}
&(\mathbf{L}\times\mathbf{p})(\mathbf{L}\times\mathbf{p}) = (\mathbf{L}\times\mathbf{p})(-\mathbf{p}\times\mathbf{L} + 2i\hbar\mathbf{p}) \\
&\quad = -(\mathbf{L}\times\mathbf{p})(\mathbf{p}\times\mathbf{L}) + 2i\hbar(\mathbf{L}\times\mathbf{p})\mathbf{p} = -(\mathbf{L}\times\mathbf{p})(\mathbf{p}\times\mathbf{L}) = p^2L^2 \qquad \text{(G.59)}
\end{aligned}
$$

and

$$
\begin{aligned}
&(\mathbf{p}\times\mathbf{L})(\mathbf{L}\times\mathbf{p}) = (-\mathbf{L}\times\mathbf{p} + 2i\hbar\mathbf{p})(\mathbf{L}\times\mathbf{p}) \\
&\quad = -(\mathbf{L}\times\mathbf{p})(\mathbf{L}\times\mathbf{p}) + 2i\hbar\mathbf{p}(\mathbf{L}\times\mathbf{p}) = -p^2L^2 + 2i\hbar\mathbf{p}(-\mathbf{p}\times\mathbf{L} + 2i\hbar\mathbf{p}) \\
&\quad = -p^2L^2 - 2i\hbar\mathbf{p}(\mathbf{p}\times\mathbf{L}) - 4\hbar^2\mathbf{p}^2 = -p^2L^2 - 4\hbar^2 p^2. \qquad \text{(G.60)}
\end{aligned}
$$

(b) The terms $\sim g$ are

$$
\begin{aligned}
&(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L})\frac{\mathbf{r}}{r} + \frac{\mathbf{r}}{r}(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}) \\
&\quad = (\mathbf{L}\times\mathbf{p})\frac{\mathbf{r}}{r} - (\mathbf{p}\times\mathbf{L})\frac{\mathbf{r}}{r} + \frac{\mathbf{r}}{r}(\mathbf{L}\times\mathbf{p}) - \frac{\mathbf{r}}{r}(\mathbf{p}\times\mathbf{L}). \qquad \text{(G.61)}
\end{aligned}
$$

Because of (G.23), it follows that

$$(\mathbf{L}\times\mathbf{p})\,\mathbf{r} = \mathbf{L}\,(\mathbf{p}\times\mathbf{r}) = -L^2. \qquad \text{(G.62)}$$

Since $\mathbf{L}$ contains only derivatives with respect to angles, we have

$$(\mathbf{L}\times\mathbf{p})\frac{\mathbf{r}}{r} = \mathbf{L}\left(\mathbf{p}\times\frac{\mathbf{r}}{r}\right) = -\frac{L^2}{r}. \qquad \text{(G.63)}$$

Furthermore, we have:

$$(\mathbf{p}\times\mathbf{L})\,\frac{\mathbf{r}}{r} = (-\mathbf{L}\times\mathbf{p} + 2i\hbar\mathbf{p})\,\frac{\mathbf{r}}{r} = \frac{L^2}{r} + 2i\hbar\mathbf{p}\frac{\mathbf{r}}{r}. \tag{G.64}$$

For the last term here, it holds that:

$$2i\hbar\mathbf{p}\frac{\mathbf{r}}{r} = 2i\hbar\frac{\hbar}{i}\nabla\frac{\mathbf{r}}{r} = 2i\hbar\frac{\hbar}{i}\left(\frac{3}{r} - \frac{\mathbf{r}\cdot\mathbf{r}}{r^3} + \frac{\mathbf{r}}{r}\nabla\right) = 2\hbar^2\frac{2}{r} + 2i\hbar\frac{\mathbf{r}}{r}\mathbf{p}. \tag{G.65}$$

Moreover, we have

$$\frac{\mathbf{r}}{r}\,(\mathbf{p}\times\mathbf{L}) = \frac{1}{r}\,(\mathbf{r}\times\mathbf{p})\,\mathbf{L} = \frac{L^2}{r} \tag{G.66}$$

and

$$\frac{\mathbf{r}}{r}\,(\mathbf{L}\times\mathbf{p}) = \frac{\mathbf{r}}{r}\,(-\mathbf{p}\times\mathbf{L} + 2i\hbar\mathbf{p}) = -\frac{L^2}{r} + 2i\hbar\frac{\mathbf{r}}{r}\mathbf{p}. \tag{G.67}$$

Hence, it follows for the terms $\sim g$:

$$\begin{aligned}
&(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L})\,\frac{\mathbf{r}}{r} + \frac{\mathbf{r}}{r}\,(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}) \\
&= -\frac{L^2}{r} - \left[\frac{L^2}{r} + 2\hbar^2\frac{2}{r} + 2i\hbar\frac{\mathbf{r}}{r}\mathbf{p}\right] + \left[-\frac{L^2}{r} + 2i\hbar\frac{\mathbf{r}}{r}\mathbf{p}\right] - \left[\frac{L^2}{r}\right] \\
&= -\frac{4L^2}{r} - \frac{4\hbar^2}{r}.
\end{aligned} \tag{G.68}$$

(c) In sum, we have:

$$\begin{aligned}
\mathbf{\Lambda}^2 &= g^2\,(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L})\,(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}) \\
&\quad + g\,(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L})\,\frac{\mathbf{r}}{r} + g\frac{\mathbf{r}}{r}\,(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}) + 1 \\
&= g^2\left\{4p^2L^2 + 4\hbar^2p^2\right\} + g\left\{-\frac{4L^2}{r} - \frac{4\hbar^2}{r}\right\} + 1 \\
&= 4g^2p^2L^2 + 4\hbar^2g^2p^2 - g\frac{4L^2}{r} - g\frac{4\hbar^2}{r} + 1 \\
&= 4\left(g^2p^2 - g\frac{1}{r}\right)L^2 + 4\hbar^2\left(g^2p^2 - g\frac{1}{r}\right) + 1.
\end{aligned} \tag{G.69}$$

With

$$\frac{p^2}{2m} - \frac{\gamma}{r} = H;\ p^2 = 2mH + 2m\frac{\gamma}{r} \tag{G.70}$$

it follows that

$$\begin{aligned}\mathbf{\Lambda}^2 &= 4\left(g^2p^2 - g\frac{1}{r}\right)\left(L^2 + \hbar^2\right) + 1\\ &= 4\left(g^2 2mH + g^2 2m\frac{\gamma}{r} - g\frac{1}{r}\right)\left(L^2 + \hbar^2\right) + 1.\end{aligned} \tag{G.71}$$

In order that the terms $\sim 1/r$ vanish, we must require:

$$g^2 2m\gamma - g = 0;\ g = \frac{1}{2m\gamma}, \tag{G.72}$$

which is in fact the case; we thus finally obtain

$$\mathbf{\Lambda}^2 = \frac{2}{m\gamma^2}H\left(L^2 + \hbar^2\right) + 1. \tag{G.73}$$

(d) Prove the two equations

$$[\mathbf{J}_1, \mathbf{J}_2] = 0;\ \mathbf{J}_1^2 = \mathbf{J}_2^2 \tag{G.74}$$

with $\mathbf{J}_1 = \frac{1}{2}(\mathbf{L} + \mathbf{R})$ and $\mathbf{J}_2 = \frac{1}{2}(\mathbf{L} - \mathbf{R})$ and $\boldsymbol{R} = \sqrt{\frac{m\gamma^2}{2|E|}}\mathbf{\Lambda}$.
Solution: We have initially

$$\begin{aligned}[\mathbf{J}_1, \mathbf{J}_2] = \mathbf{J}_1\mathbf{J}_2 - \mathbf{J}_2\mathbf{J}_1 &= \frac{1}{2}(\mathbf{L} + \mathbf{R})\frac{1}{2}(\mathbf{L} - \mathbf{R}) - \frac{1}{2}(\mathbf{L} - \mathbf{R})\frac{1}{2}(\mathbf{L} + \mathbf{R})\\ &= \frac{1}{4}\left\{\mathbf{L}^2 - \mathbf{LR} + \mathbf{RL} + \mathbf{R}^2\right\} - \frac{1}{4}\left\{\mathbf{L}^2 + \mathbf{LR} - \mathbf{RL} + \mathbf{R}^2\right\} = 0.\end{aligned} \tag{G.75}$$

For the last step, we have used $\mathbf{L}R = R\mathbf{L} = 0$.
Concerning the proof of the second equation, we note that:

$$\begin{aligned}\mathbf{J}_1^2 &= \frac{1}{4}\left(\mathbf{L}^2 + \mathbf{LR} + \mathbf{RL} + \mathbf{R}^2\right) = \frac{1}{4}\left(\mathbf{L}^2 + \mathbf{R}^2\right)\\ \mathbf{J}_2^2 &= \frac{1}{4}\left(\mathbf{L}^2 - \mathbf{LR} - \mathbf{RL} + \mathbf{R}^2\right) = \frac{1}{4}\left(\mathbf{L}^2 + \mathbf{R}^2\right).\end{aligned} \tag{G.76}$$

We have again used $\mathbf{L}\boldsymbol{R} = \boldsymbol{R}\mathbf{L} = 0$.

3. Show that $\mathbf{J}_1$ (and therefore also $\mathbf{J}_2$) is a generalized angular-momentum operator. Solution: We consider $\left[J_{1x}, J_{1y}\right]$ in more detail and infer the other relations from cyclic permutations. We have

$$\begin{aligned}\left[J_{1x}, J_{1y}\right] &= \frac{1}{2}\left(L_x + R_x\right)\frac{1}{2}\left(L_y + R_y\right) - \frac{1}{2}\left(L_y + R_y\right)\frac{1}{2}\left(L_x + R_x\right) \\ &= \frac{1}{4}\left\{L_xL_y + L_xR_y + R_xL_y + R_xR_y\right\} \\ &\quad - \frac{1}{4}\left\{L_yL_x + L_yR_x + R_yL_x + R_yR_x\right\} \\ &= \frac{1}{4}\left\{\left[L_x, L_y\right] + \left[L_x, R_y\right] + \left[R_x, L_y\right] + \left[R_x, R_y\right]\right\}. \end{aligned} \tag{G.77}$$

For the individual commutators, it holds that:

$$\begin{aligned}\left[L_x, L_y\right] &= i\hbar L_z; \quad \left[R_x, R_y\right] = i\hbar L_z \\ \left[L_x, R_y\right] &= i\hbar R_z; \quad \left[R_x, L_y\right] = i\hbar R_z,\end{aligned} \tag{G.78}$$

and therefore it follows that

$$\left[J_{1x}, J_{1y}\right] = \frac{i\hbar}{2}\left(L_z + R_z\right) = i\hbar J_{1z} \tag{G.79}$$

and the corresponding other relations follow by cyclic permutation.

Appendix H
Perturbative Calculation of the Hydrogen Atom

In this section, we want to outline the perturbation calculation for

$$H = H^{(0)} + W_{mp} + W_{ls} + W_D;\ H^{(0)} = \frac{\mathbf{p}^2}{2m} - \frac{\gamma}{r};\ \gamma = \frac{Ze^2}{4\pi\varepsilon_0} \tag{H.1}$$

in a little more detail than was given in Chap. 19. We have:

$$\begin{aligned} W_{mp} &= -\frac{\mathbf{p}^4}{8m^3c^2} \\ W_{ls} &= \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV(r)}{dr}\mathbf{l}\cdot\mathbf{s} = \frac{1}{2m^2c^2}\frac{\gamma}{r^3}\mathbf{l}\cdot\mathbf{s} \\ W_D &= \frac{\hbar^2}{8m^2c^2}\nabla^2 V(r) = \frac{\pi\hbar^2\gamma}{2m^2c^2}\delta(\mathbf{r})\,. \end{aligned} \tag{H.2}$$

With the orthonormal states $\left|n;\, j, m_j;\, l\right\rangle$, we obtain the energy corrections as

$$\left\langle n;\, j', m_j';\, l'\right| W \left|n;\, j, m_j;\, l\right\rangle. \tag{H.3}$$

For brevity we use in the following the notation

$$\langle A\rangle := \left\langle n;\, j', m_j';\, l\right| A \left|n;\, j, m_j;\, l\right\rangle. \tag{H.4}$$

H.1 Calculation of the Matrix Elements

H.1.1 Matrix Elements of W_{mp}

Because of $\frac{\mathbf{p}^2}{2m} = H^{(0)} + \frac{\gamma}{r}$, we have

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture
Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

$$\langle W_{mp}\rangle = \left\langle -\frac{\mathbf{p}^4}{8m^3c^2}\right\rangle = \left\langle -\frac{1}{2mc^2}\left(H^{(0)} + \frac{\gamma}{r}\right)^2\right\rangle$$
$$= -\frac{1}{2mc^2}\left\langle \left(H^{(0)}\right)^2 + H^{(0)}\frac{\gamma}{r} + \frac{\gamma}{r}H^{(0)} + \frac{\gamma^2}{r^2}\right\rangle. \quad \text{(H.5)}$$

Since H^0 is Hermitian, it follows that

$$\langle W_{mp}\rangle = -\frac{1}{2mc^2}\left\langle \left(E_n^{(0)}\right)^2 + 2E_n^{(0)}\frac{\gamma}{r} + \frac{\gamma^2}{r^2}\right\rangle. \quad \text{(H.6)}$$

This leads to

$$\langle W_{mp}\rangle = -\frac{1}{2mc^2}\left\{\left(E_n^{(0)}\right)^2 + 2E_n^{(0)}\gamma\left\langle\frac{1}{r}\right\rangle + \gamma^2\left\langle\frac{1}{r^2}\right\rangle\right\}\delta_{j'j}\delta_{m'_j m_j}. \quad \text{(H.7)}$$

H.1.2 Matrix Elements of W_{ls}

We have:

$$\langle W_{ls}\rangle = \left\langle\frac{1}{2m^2c^2}\frac{\gamma}{r^3}\mathbf{l}\cdot\mathbf{s}\right\rangle = \frac{\gamma}{2m^2c^2}\left\langle\frac{1}{r^3}\mathbf{l}\cdot\mathbf{s}\right\rangle. \quad \text{(H.8)}$$

Making use of

$$\mathbf{j}^2 = \mathbf{l}^2 + 2\mathbf{l}\cdot\mathbf{s} + \mathbf{s}^2 \text{ or } \mathbf{l}\cdot\mathbf{s} = \frac{1}{2}\left(\mathbf{j}^2 - \mathbf{l}^2 - \mathbf{s}^2\right), \quad \text{(H.9)}$$

we obtain:

$$\langle W_{ls}\rangle = \frac{\gamma\hbar^2}{2m^2c^2}\frac{1}{2}\left[j\,(j+1) - l\,(l+1) - \frac{3}{4}\right]\left\langle\frac{1}{r^3}\right\rangle\delta_{j'j}\delta_{m'_j m_j}. \quad \text{(H.10)}$$

For the case $l = 0$ (where the term $\left\langle\frac{1}{r^3}\right\rangle$ does not exist), we find no contribution, since then $j\,(j+1) - l\,(l+1) - \frac{3}{4} = 0$.

H.1.3 Matrix Elements of W_D

We have:

$$\langle W_D\rangle = \left\langle\frac{\pi\hbar^2\gamma}{2m^2c^2}\delta\left(\mathbf{r}\right)\right\rangle = \frac{\pi\hbar^2\gamma}{2m^2c^2}\left\langle\delta\left(\mathbf{r}\right)\right\rangle. \quad \text{(H.11)}$$

Due to the delta function, this term involves only s orbitals, because only for these is $\psi(0) \neq 0$ (because of $R_{nl} \sim r^l$ for small r). Thus, we obtain the contribution:

$$\langle W_D \rangle = \frac{\pi\hbar^2\gamma}{2m^2c^2} \left|R_{n0}(0)\right|^2 \delta_{j'j}\delta_{m'_j m_j} \tag{H.12}$$

for $l = 0$ only.

H.2 Fine Structure Corrections

Summed up: All the corrections are diagonal in j and m_j. We add them and obtain as the total correction to the energy:

$$E_n^{(1)} = \begin{bmatrix} -\frac{1}{2mc^2}\left\{\left(E_n^{(0)}\right)^2 + 2E_n^{(0)}\gamma\left\langle\frac{1}{r}\right\rangle + \gamma^2\left\langle\frac{1}{r^2}\right\rangle\right\} \\ +\frac{\gamma\hbar^2}{2m^2c^2}\frac{1}{2}\left[j(j+1) - l(l+1) - \frac{3}{4}\right]\left\langle\frac{1}{r^3}\right\rangle + \frac{\pi\hbar^2\gamma}{2m^2c^2}\left|R_{n0}(0)\right|^2\delta_{l,0} \end{bmatrix} \delta_{j'j}\delta_{m'_j m_j}. \tag{H.13}$$

Finally, we have to calculate the three matrix elements

$$\left\langle\frac{1}{r^a}\right\rangle = \left\langle n; j, m_j; l\right| \frac{1}{r^a}\left|n; j, m_j; l\right\rangle \sim \int R_{nl}^2 \frac{1}{r^a} r^2 \mathrm{d}r;\ a = 1, 2, 3. \tag{H.14}$$

We take the result from Appendix B, Vol. 2, 'Special functions':

$$\left\langle\frac{1}{r}\right\rangle = \frac{Z}{a_0 n^2};\ \left\langle\frac{1}{r^2}\right\rangle = \frac{Z^2}{a_0^2 n^3}\frac{1}{l+\frac{1}{2}};\ \left\langle\frac{1}{r^3}\right\rangle = \frac{Z^3}{a_0^3 n^3}\frac{1}{l\left(l+\frac{1}{2}\right)(l+1)}. \tag{H.15}$$

Appendix I
The Production of Entangled Photons

Entangled photons are of great importance for experiments in basic research as well as for practical applications, from quantum cryptography to quantum teleportation to quantum computing. In the following, we outline three experimental production methods. In each case, photon pairs are generated which are entangled with respect to their polarization, i.e. they are in a Bell state as $|\Psi\rangle = \frac{|hv\rangle - |vh\rangle}{\sqrt{2}}$.

I.1 Atomic Sources

In this method, appropriate atoms such as calcium or mercury are excited so that they emit two polarization-entangled photons on returning to their ground states. Calcium, for example, has in its ground state (1) two electrons in the 4s shell. By irradiation with UV light, an excited state (3, both electrons in the 4p state) is populated; it decays through a cascade via an intermediate state (2),[25] see Fig. I.1. In the first step 3→2, a photon with $\lambda = 551.3$ nm is emitted; in the second step, $2 \rightarrow 1$, a photon is emitted with $\lambda = 422.7$ nm.

It is essential that the intermediate level be degenerate with respect to the magnetic quantum number m (that is, with respect to the polarization). Assuming that the two photons are emitted along the quantization axis, the $m = 0$ state does not contribute to the optical transition. Thus, if we consider photons which are emitted in opposite directions, their state of polarization is given by

$$|\psi\rangle = \frac{|r\rangle\,|r\rangle\; +\; |l\rangle\,|l\rangle}{\sqrt{2}} = \frac{|rr\rangle\; +\; |ll\rangle}{\sqrt{2}} = \frac{|hh\rangle - |vv\rangle}{\sqrt{2}}. \tag{I.1}$$

We see that the linear polarization direction is not specified, but that the two photons are polarized parallel to each other.

[25] An electron that is in the highest level (3) cannot go directly into the ground state (1), because then angular momentum would not be conserved.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

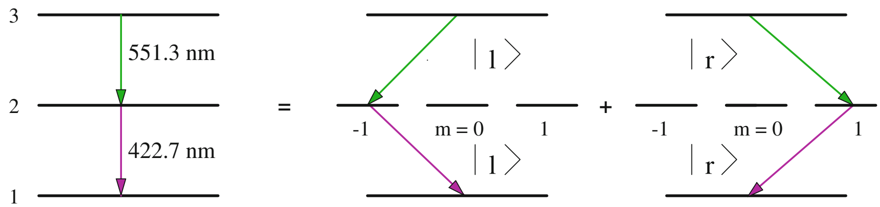


Fig. I.1 Principle of the cascade transition in calcium

This was the first method that was reliable enough to be used in experiments. However, it has serious drawbacks—among other things, the fact that the photons are not emitted in fixed directions, but rather are randomly distributed over the whole solid angle, so that the method is not very efficient.

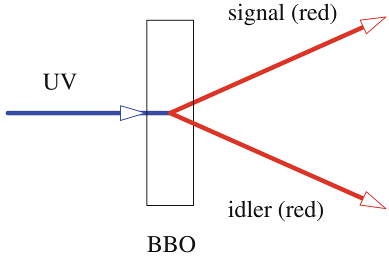
I.2 Parametric Fluorescence

In this process, one slices a photon so to speak into two daughter photons with longer wavelengths. The ‘scissors’ is an optically active (non-linear) crystal such as barium borate. It converts a UV photon into two red photons, commonly referred to as *signal* and *idler*, that are emitted in certain directions due to the conservation of momentum and energy; cf. Fig. I.2. Because of this conversion to lower photon energies, the process is also called ‘parametric down conversion’ (PDC).

Due to the polarization of the daughter photons, one can generally distinguish two types: Type I fluorescence (orthogonal polarizations), and type II fluorescence (parallel polarizations). We will hereafter consider only type II fluorescence, as only in this case can entanglement be obtained directly. In addition, we confine ourselves to the case that both daughter photons have the same energy (degenerate fluorescence).

In contrast to the two-photon cascade emission, here the two daughter photons are generated simultaneously. Above all, conservation of momentum ensures a restriction of the possible directions of emission of the signal and idler photons; namely, one can

Fig. I.2 Scheme of parametric fluorescence



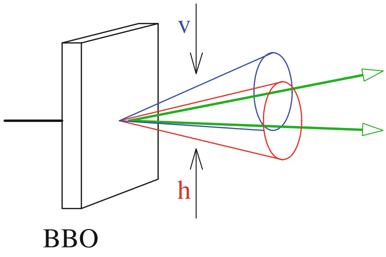


Fig. I.3 Parametric fluorescence: photons with horizontal (*bottom*) and vertical (*top*) polarization are emitted along the surfaces of two cones. A pair of photons which is emitted along the section lines of the cones is polarization-entangled

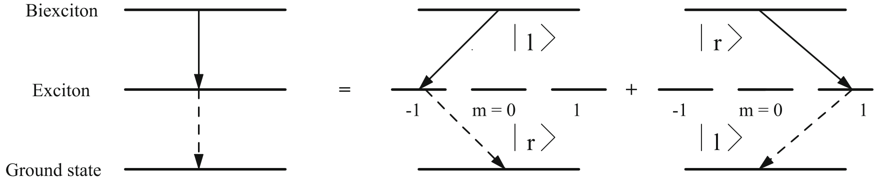


Fig. I.4 Generation of an entangled photon pair via an exciton cascade process

obtain emission of the photons along two conical surfaces. In the case of degenerate fluorescence, these emission cones have the same opening angles; see Fig. I.3.

The opening angle of the cones depends on the angle between the optical axis of the crystal and the direction of the incident light beam. In particular, one can arrange this so that the two cones intersect. If a photon is emitted in one of these intersecting directions, then the other photon is emitted in the other direction. The polarization state of this photon pair is entangled; it is given by:

$$|\psi\rangle = \frac{|hv\rangle + |vh\rangle}{\sqrt{2}}. \tag{I.2}$$

Parametric fluorescence is currently the most common method for generating entangled photons. However, the yield is limited to a few percent.

I.3 Semiconductor Sources

Here, there are several methods. One is to produce a biexciton, i.e. a state of two bound excitons.[26] Under appropriate conditions, the biexciton decays in a cascade process that is quite similar to the atomic decay described above; see Fig. I.4.[27]

[26] An exciton is a bound state between an electron and a hole in a semiconductor or insulator.

[27] In other words, the photon pair is produced by the radiative decay of two electron-hole pairs.

This results in an entangled pair of photons in the state

$$|\psi\rangle = \frac{|hh\rangle + |vv\rangle}{\sqrt{2}}. \tag{I.3}$$

There is no fundamental limitation of the yield here, and the mechanism is quite effective. However, the photon pairs are emitted isotropically, in all directions.

I.4 Concluding Remarks

Given the importance of entangled photons, it is understandable that researchers are actively working on reducing the limitations of the procedures described. With parametric fluorescence, one can e.g. enhance the entrance channel by 'amplifying' the UV pulses in a resonator. Thereby, ultrashort light pulses in very rapid succession are produced, with which one can obtain much higher emission rates of entangled photons, and even more pairs which are entangled with one another.[28]

With semiconductor sources, for example, a method was developed that allows for effective collection of the photon pairs generated; keyword 'photonic molecule'. Using this method, an efficiency of 80–90 % should be achievable, i.e. 8–9 photon pairs per 10 excitation pulses.[29] By means of quantum points it is possible to produce entangled photons so to say off the production line.[30]

Finally, we note that apart from photons/polarization, one can of course entangle other quantum objects/properties. An example is the spin entanglement of atoms. Here, a suitable diatomic molecule with zero spin is excited so highly that it dissociates, i.e. it decays into two atoms, each with spin 1/2. This atomic pair is spin entangled, where the entanglement of course refers to the z components of the spins.

---

[28] R. Krischek et al., 'Ultraviolet enhancement cavity for ultrafast nonlinear optics and high-rate multiphoton entanglement experiments.' *Nature Photonics* 4, 170–173 (2010).

[29] A. Dousse et al., 'Ultrabright source of entangled photon pairs.' *Nature* 466, 217 (2010).

[30] See I. Schwartz et al., Deterministic generation of a cluster state of entangled photons, *Science*, online: 8. September 2016; http://dx.doi.org/10.1126/science.aah4758.

Appendix J
The Hardy Experiment

J.1 The Experiment

Interaction-free measurement and quantum entanglement—these two elements of quantum mechanics are combined in Hardy's experiment,[31] shown schematically in Fig. J.1. The point is, *inter alia*, that we can entangle two quantum objects in an 'interaction-free' manner[32] here.

An electron-positron pair (or another particle-antiparticle pair) enters two Mach–Zehnder interferometers[33] (MZI) at the same time, the positron into the upper and the electron into the lower MZI. We denote the properties of the positron and the electron in the following as 1 and 2. If positron and electron meet at the point WW, they annihilate with certainty, $e^+e^- \rightarrow 2\gamma$.

In both MZI's, the first beam splitter BS and the mirror M are fixed. In each case, the second beam splitter, i.e. BS1 or BS2, can be moved out of/into the beam path. At WW, the arms of the two MZI's cross. Here, by an appropriate arrangement of the arms, we can choose either that the electron and the positron (a) do not meet and thus are not annihilated; or (b) do meet and thus are annihilated.

We consider first the situation in which the destructive interaction is switched on and both beam splitters BSi are in place. As described in Chap. 6, Vol. 1, each MZI is adjusted in such a way that a quantum object entering horizontally or vertically is always detected at the horizontal or vertical detector (the 'bright' detectors). If the other ('dark') detectors register a quantum, we know that some disturbing object is

[31] Lucien Hardy, 'Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories.' *Phys. Rev. Lett.* 68, 2981 (1992).

[32] For the limitations on the term 'interaction-free', see Chap. 6, Vol. 1.

[33] Of course, the components are made in such a way that they in fact perform as beam splitters and mirrors for the electrons and positrons. By the way, recent experiments were carried out with two photons, where the pair annihilation is replaced by a process of destructive interference. Here, a special type of quantum measurement was used (called *weak measurement*); see J.S. Lundeen & A.M. Steinberg, 'Experimental joint weak measurement on a photon pair as a probe of Hardy's paradox', arXiv:0810.4229 (2008).

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

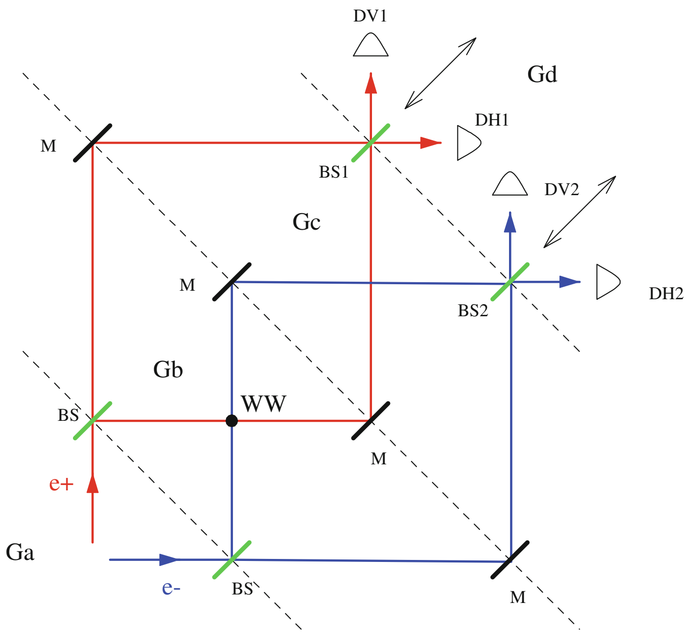


Fig. J.1 Hardy's experiment. The beam splitters BS and mirrors M are fixed, the beam splitters BSi can be moved in/out of the ray path. Electrons (*blue*) and positrons (*red*) are detected in the detectors DVi and DHi

in the beam path. That means in Hardy's experiment that a click from a dark detector indicates that the quantum object must have 'probed' the arm with WW. When both dark detectors register a count, i.e. DH1 and DV2, the electron and the positron must have simultaneously 'probed' the WW point and thus should have annihilated with certainty. But that would contradict the result that they were both detected by the two dark detectors, which in fact is observed in a certain percentage of cases. From a classical point of view, this is a contradiction, known as Hardy's paradox.

We now describe the reactions of the detectors as a function of whether we turn the annihilation interaction at WW on or off, and whether the beam splitters BS1, 2 are in the path or not. We give the calculation further below, and just anticipate the result here.

For simplicity, we consider for the moment only the case that the two movable detectors are in the path of the setup (J.15). The probabilities for the activation of the two detectors are given in Table J.1:

For inactivated annihilation ($\alpha = 1$), we find again the result of Chap. 6, Vol. 1, indicating that the setup reproduces the initial state. Thus, detectors (DV1, DH2) are 'bright' and always respond at the same time.

For activated annihilation ($\alpha = 0$), a quarter of the positron-electron pairs annihilates, and three quarters can be detected in the detectors. The pairs landing in (DH1, DH2), (DH1, DV2) and (DV1, DV2) cannot have interacted directly by means of

Table J.1 Probabilities for the simultaneous activation of two detectors

Detectors	$\alpha = 1$	$\alpha = 0$
DH1,DH2	0	1/16
DH1,DV2	0	1/16
DV1,DH2	1	9/16
DV1,DV2	0	1/16

For $\alpha = 0$ or 1, the annihilation process is activated or deactivated

the pair annihilation, as they would then have been annihilated; but they must have 'realized' the pair annihilation, since otherwise neither one nor two 'dark' detectors would have responded. In particular, in one sixteenth of the cases, both dark detectors respond. This means that both the electron and the positron have 'realized' that a disturbing object is in their beam paths—and this, of course, could be only the other partner. Consequently, e^+ and e^- must have been at the same point (WW) in some way, without having destroyed each other.

J.2 Calculation of the Probabilities

For the calculation, we adopt the notation of Chap. 6, Vol. 1, namely that $|H\rangle$ and $|V\rangle$ denote horizontal and vertical propagation channels. Moreover, we derived there for the operators T (beam splitter) and S (mirror) the expressions:

$$\begin{aligned} T &= \tfrac{1+i}{2}\left[1 + i\,|H\rangle\langle V| + i\,|V\rangle\langle H|\right] \\ S &= -\,|H\rangle\langle V| - |V\rangle\langle H|\,. \end{aligned} \tag{J.1}$$

The annihilation interaction at point WW can be effective only if in the first MZI, a horizontal state is present, and in the second MZI, a vertical state. We can, therefore, represent this by an operator W_α with the following properties:

$$\begin{aligned} &W_\alpha\,|H_1H_2\rangle = |H_1H_2\rangle\,;\;\; W_\alpha\,|H_1V_2\rangle = \alpha\,|H_1V_2\rangle \\ &W_\alpha\,|V_1H_2\rangle = |V_1H_2\rangle\,;\;\; W_\alpha\,|V_1V_2\rangle = |V_1V_2\rangle \\ &\alpha = \begin{cases} 1 \\ 0 \end{cases} \text{for } \begin{matrix}\text{inactivated} \\ \text{activated}\end{matrix} \text{ annihilation interaction.} \end{aligned} \tag{J.2}$$

Since $\{|H\rangle, |V\rangle\}$ is a CONS, we can write

$$\begin{aligned} W_\alpha &= |H_1H_2\rangle\langle H_1H_2| + \alpha\,|H_1V_2\rangle\langle H_1V_2| + |V_1H_2\rangle\langle V_1H_2| + |V_1V_2\rangle\langle V_1V_2| \\ &= 1 + (\alpha - 1)\,|H_1V_2\rangle\langle H_1V_2|\,. \end{aligned} \tag{J.3}$$

This enables us to formulate the effect of the setup shown in Fig. J.1. We employ here the method used in Chap. 6, Vol. 1, namely dividing the setup into four main regions: In the first region Ga, we have the incoming state $|z_a\rangle$; in Gb the state $|z_b\rangle = T_1T_2|z_a\rangle$ transformed by the beam splitter, which possibly is modified by the annihilation interaction W_α.[34] This state $|z_b\rangle = W_\alpha T_1T_2|z_a\rangle$ is reflected by the mirrors, so that in Gc, we have $|z_c\rangle = S_1S_2W_\alpha T_1T_2|z_1\rangle$. Finally, in Gd, following the beam splitters BS1, 2, we have the final state $|z_d\rangle$. We denote the operator representing the double MZI by M and distinguish three cases:

$$\begin{array}{ll} \text{1. no beam splitters BS1,2 inserted} & M(0) = S_1S_2W_\alpha T_1T_2 \\ \text{2. one beam splitter BS1,2 inserted} & M(1,i) = T_iS_1S_2W_\alpha T_1T_2;\; i=1,2 \\ \text{3. both beam splitters BS1,2 inserted} & M(2) = T_1T_2S_1S_2W_\alpha T_1T_2. \end{array} \tag{J.4}$$

In the following, we calculate M in the matrix representation. We start from

$$|H\rangle \cong \begin{pmatrix} 1 \\ 0 \end{pmatrix};\; |V\rangle \cong \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{J.5}$$

and obtain (cf. Chap. 6, Vol. 1):

$$T \cong \frac{1+i}{2}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix};\; S \cong -\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{J.6}$$

For the direct product of two beam splitters or mirrors, it follows that:

$$T_1T_2 \cong \frac{i}{2}\begin{pmatrix} 1 & i & i & -1 \\ i & 1 & -1 & i \\ i & -1 & 1 & i \\ -1 & i & i & 1 \end{pmatrix};\; S_1S_2 \cong \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{J.7}$$

If we use only one beam splitter, the matrices (E is the 2×2 unit matrix) are given by:

$$T_1E_2 = \frac{(1+i)}{2}\begin{pmatrix} 1 & 0 & i & 0 \\ 0 & 1 & 0 & i \\ i & 0 & 1 & 0 \\ 0 & i & 0 & 1 \end{pmatrix};\; E_1T_2 = \frac{(1+i)}{2}\begin{pmatrix} 1 & i & 0 & 0 \\ i & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & i & 1 \end{pmatrix}. \tag{J.8}$$

The annihilation interaction is then represented by

[34] We leave off the explicit specification of the direct product, and write simply T_1T_2 instead of $T_1 \otimes T_2$.

$$W_\alpha \cong \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{J.9}$$

Now we can calculate the operators M. We have for $M\ (0)$:

$$M\ (0) \cong \frac{i}{2} \begin{pmatrix} -1 & i & i & 1 \\ i & -1 & 1 & i \\ \alpha i & \alpha & -\alpha & \alpha i \\ 1 & i & i & -1 \end{pmatrix}, \tag{J.10}$$

and for $M\ (1, i) = T_i M\ (0)$:

$$M\ (1, 1) \cong \frac{i}{2} \frac{(1 + i)}{2} \begin{pmatrix} -1 - a & i\ (1 + a) & i\ (1 - a) & 1 - a \\ 2i & -2 & 0 & 0 \\ -i\ (1 - a) & -1 + a & -1 - a & i\ (1 + a) \\ 0 & 0 & 2i & -2 \end{pmatrix}, \tag{J.11}$$

as well as

$$M\ (1, 2) \cong \frac{i}{2} \frac{(1 + i)}{2} \begin{pmatrix} -2 & 0 & 2i & 0 \\ 0 & -2 & 0 & 2i \\ i\ (\alpha + 1) & \alpha - 1 & -\alpha - 1 & i\ (\alpha - 1) \\ -\alpha + 1 & i\ (\alpha + 1) & i\ (-\alpha + 1) & -\alpha - 1 \end{pmatrix}. \tag{J.12}$$

For $M\ (2)$, it follows that:

$$M\ (2) \cong -\frac{1}{4} \begin{pmatrix} -3 - \alpha & i\ (\alpha - 1) & i\ (1 - \alpha) & 1 - \alpha \\ i\ (\alpha - 1) & -3 - \alpha & \alpha - 1 & i\ (1 - \alpha) \\ i\ (1 - \alpha) & \alpha - 1 & -3 - \alpha & i\ (\alpha - 1) \\ 1 - \alpha & i\ (\alpha - 1) & i\ (1 - \alpha) & -3 - \alpha \end{pmatrix}. \tag{J.13}$$

We now specialize, as shown in Fig. J.1, to the non-entangled initial state:

$$|v_1 h_2\rangle \cong \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \tag{J.14}$$

and obtain as final states $|z_d\rangle$:

$$|z_d(0)\rangle = \frac{i}{2}\begin{pmatrix} i \\ 1 \\ -\alpha \\ i \end{pmatrix}; \quad |z_d(1,1)\rangle = \frac{i}{2}\frac{(1+i)}{2}\begin{pmatrix} i(1-\alpha) \\ 0 \\ -1-\alpha \\ 2i \end{pmatrix}$$
$$|z_d(1,2)\rangle = \frac{i}{2}\frac{(1+i)}{2}\begin{pmatrix} 2i \\ 0 \\ -1-\alpha \\ i(1-\alpha) \end{pmatrix}; \quad |z_d(2)\rangle = \frac{1}{4}\begin{pmatrix} i(\alpha-1) \\ 1-\alpha \\ 3+\alpha \\ i(\alpha-1) \end{pmatrix}. \tag{J.15}$$

We wish now to discuss these expressions briefly. We consider first the detection probabilities and then the entanglement status of the e^+e^- pairs detected.

Detection Probabilities

If the final state is given e.g. as $|h_1h_2\rangle$, then both detectors DH1 and DH2 are activated. As usual, the detection probabilities are given by the sum of the squared values of the corresponding coefficients in (J.15). The results are collected in Table J.2.

We first leave the pair annihilation switched off ($\alpha = 1$). We see a uniform distribution of the results if both beam splitters are withdrawn; if both are in place, however, the setup reproduces the initial state $|v_1h_2\rangle$.

Now we let the pair annihilation act ($\alpha = 0$). In this case, one quarter of the positron-electron pairs will always annihilate, regardless of whether there are beam splitters in the beam path or not. The remaining 75 % of the pairs, which are detected by the detectors, cannot have interacted *directly* by pair annihilation, as they would have been annihilated in that case, and not detected.

We repeat our discussion on the configuration with two beam splitters BSi inserted (i.e. the final state is $|z_d(2)\rangle$). Without pair annihilation (i.e. for $\alpha = 1$), only the detector pair DV1 and DH2 responds. Thus, if a 'dark' detector responds we know that some 'obstacle' must have been in the path; but this can only be the pair annihilation at WW, i.e. $\alpha = 0$. Hence, if *both* dark detectors DH1 and DV2 respond, both the electron and the positron must have 'remarked' that a disturbing object is in the beam path, i.e. the other partner. Consequently, e^+ and e^- must have met somehow at WW without having annihilated each other—Hardy's paradox.

Table J.2 Probabilities for the simultaneous response of two detectors

Detectors	$\lvert z_d(0)\rangle$	$\lvert z_d(1,1)\rangle$	$\lvert z_d(1,2)\rangle$	$\lvert z_d(2)\rangle$
$\lvert h_1h_2\rangle \hat{=}$ (DH1, DH2)	$\frac{1}{4}$	$\frac{(1-\alpha)^2}{8}$	$\frac{1}{2}$	$\frac{(1-\alpha)^2}{16}$
$\lvert h_1v_2\rangle \hat{=}$ (DH1, DV2)	$\frac{1}{4}$	0	0	$\frac{(1-\alpha)^2}{16}$
$\lvert v_1h_2\rangle \hat{=}$ (DV1, DH2)	$\frac{\alpha^2}{4}$	$\frac{(1+\alpha)^2}{8}$	$\frac{(1+\alpha)^2}{8}$	$\frac{(3+\alpha)^2}{16}$
$\lvert v_1v_2\rangle \hat{=}$ (DV1, DV2)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{(1-\alpha)^2}{8}$	$\frac{(1-\alpha)^2}{16}$

For $\alpha = 0$ or 1, the annihilation process is switched on or off

Entanglement

We first recall the simple criterion, formulated in Chap. 20, as to whether a four-dimensional vector (as the direct product of two 2-dimensional vectors) represents an entangled state: It must fulfill $c_1c_4 \neq c_2c_3$, where c_i are the four components of the corresponding vector.

Applying this criterion to the states (J.15), we obtain Table J.3. As we can read off immediately, the final states are always factorizable for $\alpha = 1$ (no annihilation), i.e. they are not entangled. However, they are *always* entangled for activated pair annihilation ($\alpha = 0$), independently of whether BS1 and/or BS2 are in the beam path or not.

This is interesting because the initial state (J.14) is not entangled. Hence, we have the result for activated pair annihilation that a non-entangled state becomes entangled, and this in an interaction-free manner, in the sense that the electron and the positron cannot have interacted directly via pair annihilation.

By the way, the converse is not true: If we send an entangled initial state through the setup with two beam splitters BSi, it will not be dis-entangled, but rather it remains entangled. An initial state of the form $\sim |h_1h_2\rangle + |v_1v_2\rangle$ is for example converted into a final state of the same form.

Table J.3 Entanglement criterion for the final states (J.15)

	Without BSi's	With BS1	With BS2	With both BSi's
$c_1c_4 - c_2c_3$	$i^2 + \alpha$	$2i^2(1-\alpha)$	$2i^2(1-\alpha)$	$i^2(\alpha-1) - (1-\alpha)(3-\alpha)$

Appendix K
Set-Theoretical Derivation of the Bell Inequality

Bell's inequality, considered in Chap. 20, may also be derived using the calculus of set theory.

Given a total set U, from which we single out three subsets a, b and c; we label the respective (absolute) complements by $\neg a$, $\neg b$ and $\neg c$. We can combine and intersect these sets; we can form e.g. $a \cap b$.

We start with the equation

$$a \cap b = (a \cap b) \cap (c \cup \neg c) . \tag{K.1}$$

With the distributive property of sets:

$$A \cap (B \cap C) = (A \cap B) \cup (A \cap C) , \tag{K.2}$$

it follows for $A = a \cap b$, $B = c$ and $C = \neg c$ that:

$$a \cap b = (a \cap b) \cap (c \cap \neg c) = ((a \cap b) \cap c) \cup ((a \cap b) \cap \neg c) . \tag{K.3}$$

The relations

$$(a \cap b) \cap c \subseteq a \cap c; \;\; (a \cap b) \cap \neg c \subseteq b \cap \neg c \tag{K.4}$$

lead to the inequality

$$a \cap b \subseteq (a \cap c) \cup (b \cap \neg c) . \tag{K.5}$$

For the numbers (occurrence frequencies) of the elements, it holds correspondingly that

$$n(a, b) \leq n(a, c) \; + \; n(b, \neg c). \tag{K.6}$$

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

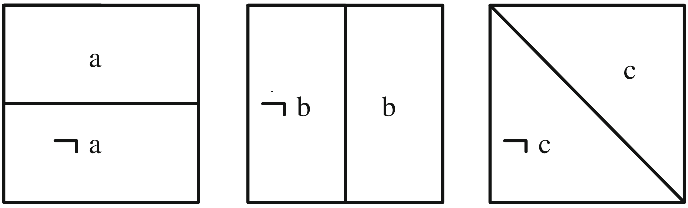


Fig. K.1 Example of the partition of a set according to the properties a, b and c

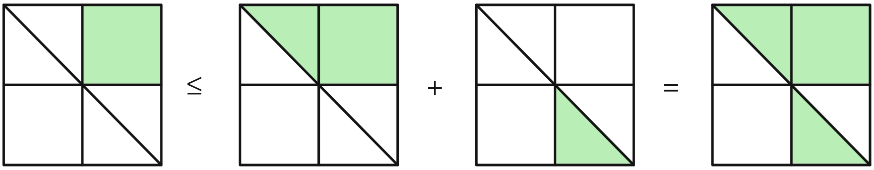


Fig. K.2 Graphical representation of inequality (K.6) $n(a, b) \leq n(a, c) + n(b, \neg c)$ for the partition as shown in Fig. K.1

We illustrate this inequality by means of the example in Fig. K.1. Then we can represent the inequality (K.6) graphically as shown in Fig. K.2.

Appendix L
The Special Galilei Transformation

Restricting ourselves to non-relativistic conditions, we consider here inertial frames related by a Galilei transformation[35] $\mathbf{r}' = \mathbf{r} + \mathbf{v}t$ with constant velocity $\mathbf{v}$, see Fig. L.1. Unlike rotations and translations, the kinetic energy is not invariant in this transformation. Nevertheless, because we are dealing with inertial frames, the form of the laws of physics, specifically the SEq, must stay the same under the transformation (shape invariance).[36] This is the case, as we show explicitly in the following. To avoid ambiguities, we denote the position and momentum operators by $\mathbf{X}$ and $\mathbf{P}$, and their eigenvalues by $\mathbf{r}$ or $\mathbf{x}$ and $\mathbf{p}$.

L.1 Special Galilei Transformation

L.1.1 Abstract Notation

We have two inertial frames S and S' which are related by $\mathbf{r}' = \mathbf{r} + \mathbf{v}t$. Since the transformation is continuous, it is represented by a unitary operator

$$U(\mathbf{v}) = e^{-i\mathbf{v}\mathbf{G}/\hbar} \tag{L.1}$$

with a Hermitian[37] operator $\mathbf{G}$. To determine $\mathbf{G}$, we translate the classical relations

$$\mathbf{r}' = \mathbf{r} + t\mathbf{v};\ \mathbf{p}' = \mathbf{p} + m\mathbf{v} \tag{L.2}$$

[35] We perform here an active coordinate transformation (the system is shifted, the coordinate system remains); it is also called a *boost*. In a passive transformation, the system remains, while the coordinate system is shifted.

[36] Otherwise, the inertial systems would not be equivalent.

[37] The minus sign in the exponent is purely conventional; $\mathbf{G}$ as in ‘Galileo’.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

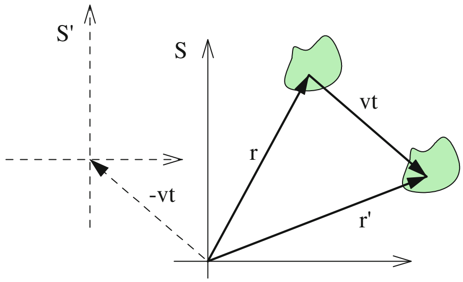


Fig. L.1 Galilei transformation

by means of

$$\mathbf{X}' = U^{-1}(\mathbf{v})\mathbf{X}U(\mathbf{v}) \overset{!}{=} \mathbf{X} + t\mathbf{v};\ \mathbf{P}' = U^{-1}(\mathbf{v})\mathbf{P}U(\mathbf{v}) \overset{!}{=} \mathbf{P} + m\mathbf{v} \tag{L.3}$$

into quantum mechanics. For sufficiently small, but otherwise arbitrary **v**, we obtain for **X** (with the infinitesimal approximation $e^{-i\mathbf{vG}/\hbar} \approx 1 - i\mathbf{vG}/\hbar$) the following expression:

$$\begin{aligned} (1 + i(\mathbf{vG})/\hbar)\,\mathbf{X}(1 - i(\mathbf{vG})/\hbar) &= \mathbf{X} + t\mathbf{v}, \\ \text{or } \mathbf{X} + i(\mathbf{vG})/\hbar\mathbf{X} - \mathbf{X}i(\mathbf{vG})/\hbar &= \mathbf{X} + t\mathbf{v}, \\ \text{or } \frac{i}{\hbar}\left((\mathbf{vG})\,\mathbf{X} - \mathbf{X}(\mathbf{vG})\right) &= t\mathbf{v}, \end{aligned} \tag{L.4}$$

and analogously for **P**. Since the last equation must hold for any velocities, it follows that

$$\frac{i}{\hbar}\left(G_i X_j - X_j G_i\right) = t\delta_{ij};\ \text{analogously}\ \frac{i}{\hbar}\left(G_i P_j - P_j G_i\right) = m\delta_{ij}. \tag{L.5}$$

To specify the exact form of $\mathbf{G} = \mathbf{G}(\mathbf{X}, \mathbf{P})$, we use a result from Chap. 21, namely

$$\left[X_i, g(P_j)\right] = i\hbar\frac{\partial g(P_j)}{\partial P_j}\delta_{ij};\ \left[P_i, f(X_j)\right] = -i\hbar\frac{\partial f(X_j)}{\partial X_j}\delta_{ij}. \tag{L.6}$$

It follows that

$$\frac{\partial G_i}{\partial P_j} = t\delta_{ij};\ \frac{\partial G_i}{\partial X_j} = -m\delta_{ij} \tag{L.7}$$

and we can take as solution

$$G_i = tP_i - mX_i\ \text{ or }\ \mathbf{G} = t\mathbf{P} - m\mathbf{X}. \tag{L.8}$$

Hence, the unitary transformation reads

$$U(\mathbf{v}) = e^{-i\mathbf{vG}/\hbar} = e^{-i\mathbf{v}(t\mathbf{P} - m\mathbf{X})/\hbar}. \tag{L.9}$$

One has to be a little careful with this operator, as the two operators $\boldsymbol{X}$ and $\boldsymbol{P}$ appearing in the exponent do not commute, $\left[X_i, P_j\right] = i\hbar\delta_{ij}$. Here, we will exploit a transformation[38] which we proved in Chap. 13, Vol. 1: Two operators A and B, both of which commute with their commutator $[A, B]$, satisfy

$$e^A e^B = e^{A+B} e^{\frac{1}{2}[A,B]}. \tag{L.10}$$

It follows (see exercises) that:

$$U(\mathbf{v}) = e^{-i\mathbf{v}(t\mathbf{P}-m\mathbf{X})/\hbar} = \begin{cases} e^{im\mathbf{vX}/\hbar} e^{-i\mathbf{vP}t/\hbar} e^{-i\frac{mv^2}{2\hbar}t} \\ e^{-i\mathbf{vP}t/\hbar} e^{im\mathbf{vX}/\hbar} e^{i\frac{mv^2}{2\hbar}t} \end{cases} \tag{L.11}$$

How are the states in S and S' related? We have

$$\left|\psi'(t)\right\rangle = U(\mathbf{v})\left|\psi(t)\right\rangle. \tag{L.12}$$

In the position representation, it follows that

$$\langle\mathbf{q}\left|\psi'(t)\right\rangle = \langle\mathbf{q}|\, U(\mathbf{v})\left|\psi(t)\right\rangle. \tag{L.13}$$

Because of $\langle\mathbf{q}|\, U(\mathbf{v}) = \left(U^\dagger(\mathbf{v})\,|\mathbf{q}\rangle\right)^\dagger$, we consider first $U^\dagger(\mathbf{v})\,|\mathbf{q}\rangle$[39]:

$$\begin{aligned} U^\dagger(\mathbf{v})\,|\mathbf{q}\rangle &= e^{-i\frac{mv^2}{2\hbar}t} e^{-im\mathbf{vX}/\hbar} e^{i\mathbf{vP}t/\hbar}\,|\mathbf{q}\rangle \\ &= e^{-i\frac{mv^2}{2\hbar}t} e^{-im\mathbf{vX}/\hbar}\,|\mathbf{q}-\mathbf{v}t\rangle = e^{-i\frac{mv^2}{2\hbar}t} e^{-im\mathbf{v}(\mathbf{q}-\mathbf{v}t)/\hbar}\,|\mathbf{q}-\mathbf{v}t\rangle. \end{aligned} \tag{L.14}$$

It follows that:

$$\langle\mathbf{q}\left|\psi'(t)\right\rangle = \langle\mathbf{q}|\, U(\mathbf{v})\left|\psi(t)\right\rangle = e^{i\frac{mv^2}{2\hbar}t} e^{im\mathbf{v}(\mathbf{q}-\mathbf{v}t)/\hbar}\,\langle\mathbf{q}-\mathbf{v}t\left|\psi(t)\right\rangle \tag{L.15}$$

or

$$\psi'(\mathbf{q}, t) = e^{-i\frac{mv^2}{2\hbar}t} e^{im\mathbf{vq}/\hbar}\psi(\mathbf{q}-\mathbf{v}t, t). \tag{L.16}$$

Finally, we set $\mathbf{q} = \mathbf{r}' = \mathbf{r} + \mathbf{v}t$ and obtain

$$\psi'\left(\mathbf{r}', t\right) = e^{-i\frac{mv^2}{2\hbar}t} e^{im\mathbf{vr}'/\hbar}\psi(\mathbf{r}, t) = e^{i\frac{mv^2}{2\hbar}t} e^{im\mathbf{vr}/\hbar}\psi(\mathbf{r}, t). \tag{L.17}$$

[38] The Baker–Campbell–Hausdorff formula, see the exercises for Chap. 13, Vol. 1.

[39] Note the sign: $e^{\pm i\mathbf{ap}/\hbar}\,|\mathbf{r}\rangle = |\mathbf{r}\mp\mathbf{a}\rangle$.

L.1.2 Position Representation

For comparison with the abstract approach of the last paragraph, we consider the effect of the transformation $\mathbf{r}' = \mathbf{r} + \mathbf{v}t$ by starting directly from the position representation. The SEq reads

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) = \left[-\frac{\hbar^2}{2m}\nabla_{\mathbf{r}}^2 + V(\mathbf{r},t)\right]\psi(\mathbf{r},t). \tag{L.18}$$

The potential may be time dependent. To distinguish between the derivatives in the two inertial systems, we use subscript indices in the derivative operators, e.g. $\nabla_{\mathbf{r}'} = \left(\frac{\partial}{\partial x'}, \frac{\partial}{\partial y'}, \frac{\partial}{\partial z'}\right)$. For the wavefunction and spatial derivative with $\mathbf{r}' = \mathbf{r} + t\mathbf{v}$, it clearly holds that

$$\psi(\mathbf{r},t) = \psi\left(\mathbf{r}' - \mathbf{v}t, t\right);\ \nabla_{\mathbf{r}}\psi(\mathbf{r},t) = \nabla_{\mathbf{r}'}\psi\left(\mathbf{r}' - \mathbf{v}t, t\right). \tag{L.19}$$

The time derivatives in the two reference systems are related by[40]:

$$\frac{\partial}{\partial t}\psi(\mathbf{r},t) = \frac{\partial}{\partial t}\psi\left(\mathbf{r}' - \mathbf{v}t, t\right) + \mathbf{v}\cdot\nabla_{\mathbf{r}'}\psi\left(\mathbf{r}' - \mathbf{v}t, t\right). \tag{L.20}$$

We insert all the intermediate results into the SEq (L.18) and obtain

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\psi\left(\mathbf{r}' - \mathbf{v}t, t\right) &+ i\hbar\mathbf{v}\cdot\nabla_{\mathbf{r}'}\psi\left(\mathbf{r}' - \mathbf{v}t, t\right) \\ &= \left[-\frac{\hbar^2}{2m}\nabla_{\mathbf{r}'}^2 + V\left(\mathbf{r}' - \mathbf{v}t, t\right)\right]\psi\left(\mathbf{r}' - \mathbf{v}t, t\right). \end{aligned} \tag{L.21}$$

Although this representation is correct, the occurrence of the argument $\mathbf{r}' - \mathbf{v}t$ and of the term $\nabla_{\mathbf{r}'}\psi$ on the left-hand side is annoying. A remedy is provided by the unitary transformation (see exercises):

$$\psi(\mathbf{r},t) = \psi\left(\mathbf{r}' - \mathbf{v}t, t\right) = e^{i\left(-m\mathbf{v}\mathbf{r}' + mv^2t/2\right)/\hbar}\psi'\left(\mathbf{r}', t\right). \tag{L.22}$$

It leads to

$$i\hbar\frac{\partial}{\partial t}\psi'\left(\mathbf{r}', t\right) = \left[-\frac{\hbar^2}{2m}\nabla_{\mathbf{r}'}^2 + V'\left(\mathbf{r}', t\right)\right]\psi'\left(\mathbf{r}', t\right);\ V'\left(\mathbf{r}', t\right) := V\left(\mathbf{r}' - \mathbf{v}t, t\right), \tag{L.23}$$

i.e. the familiar form of the SEq, however with a modified potential. Only the free SEq is completely invariant under Galilei transformations, and not just form invariant.

Inverting the unitary transformation (L.22) yields

[40] On the left side, the variable $\mathbf{r}$ is fixed in the derivative with respect to t. Therefore, also $\mathbf{r}' - \mathbf{v}t$ must be constant, and thus the term $-\mathbf{v}\nabla_{\mathbf{r}'}\psi\left(\mathbf{r}' - \mathbf{v}t, t\right)$ must be subtracted on the right side.

$$\psi'\left(\mathbf{r}', t\right) = e^{i\left(m\mathbf{v}\mathbf{r}' - mv^2t/2\right)/\hbar}\psi\left(\mathbf{r}, t\right), \tag{L.24}$$

i.e. the same result as (L.17). As an example, the transformation of a plane wave is considered in the exercises.

L.1.3 Several Quantum Objects

We start with a closed system of two quantum objects. The interaction depends only on the relative distance **r**; external interactions do not exist. We can thus formulate the SEq using the relative coordinate **r** and the center-of-mass coordinate **R** (see Appendix E, Vol. 2):

$$i\hbar\partial_t\Psi\left(\mathbf{R}, \mathbf{r}, t\right) = \left[-\frac{\hbar^2}{2M}\nabla^2_{\mathbf{R}} - \frac{\hbar^2}{2\mu}\nabla^2_{\mathbf{r}} + V\left(\mathbf{r}\right)\right]\Psi\left(\mathbf{R}, \mathbf{r}, t\right) \tag{L.25}$$

where M is the total mass and μ the reduced mass.

A Galilei boost leaves the relative coordinate unchanged, while for the center-of-mass coordinate, we see that

$$\mathbf{R}' = \mathbf{R} + \mathbf{v}t. \tag{L.26}$$

In the system S', $\Psi\left(\mathbf{R}, \mathbf{r}, t\right)$ becomes $\Psi\left(\mathbf{R}' - \mathbf{v}t, \mathbf{r}, t\right)$, and we obtain with the transformation

$$\Psi\left(\mathbf{R}' - \mathbf{v}t, \mathbf{r}, t\right) = e^{i\left(-M\mathbf{v}\mathbf{R}' + Mv^2t/2\right)/\hbar}\Psi'\left(\mathbf{R}', \mathbf{r}, t\right) \tag{L.27}$$

the new SEq

$$i\hbar\partial_t\Psi'\left(\mathbf{R}' - \mathbf{v}t, \mathbf{r}, t\right) = \left[-\frac{\hbar^2}{2M}\nabla^2_{\mathbf{R}'} - \frac{\hbar^2}{2\mu}\nabla^2_{\mathbf{r}} + V\left(\mathbf{r}\right)\right]\Psi'\left(\mathbf{R}' - \mathbf{v}t, \mathbf{r}, t\right). \tag{L.28}$$

For N quantum objects, the considerations proceed analogously. The transformation for the position and momentum operators of the quantum object with index n is

$$U^{-1}(\mathbf{v})\mathbf{X}_n U(\mathbf{v}) = \mathbf{X}_n + t\mathbf{v};\ U^{-1}(\mathbf{v})\mathbf{P}_n U(\mathbf{v}) = \mathbf{P}_n + m\mathbf{v}. \tag{L.29}$$

Since the operators for different quantum objects commute, we have

$$U(\mathbf{v}) = \prod_n e^{-i\mathbf{v}(t\mathbf{P}_n - m\mathbf{X}_n)/\hbar} \tag{L.30}$$

and it follows that

$$\Psi'\left(\mathbf{R}', \mathbf{r}_1, \ldots, \mathbf{r}_N, t\right) = e^{iM\mathbf{v}\mathbf{R}'/\hbar} e^{-i\frac{Mv^2}{2\hbar}t} \Psi\left(\mathbf{R}', \mathbf{r}_1 - \mathbf{v}t, \ldots, \mathbf{r}_N - \mathbf{v}t, t\right). \quad \text{(L.31)}$$

L.2 The Special Galilei Transformation and Kinetic Energy

In this section, we consider the most general form of a Hamiltonian which is compatible with the special Galilei transformation. For this purpose, we start from the transformation (L.9), where we confine ourselves to $t = 0$ for the sake of simplicity.

We recapitulate briefly: The requirement

$$U^{-1}(\mathbf{v})\mathbf{P}U(\mathbf{v}) \stackrel{!}{=} \mathbf{P} + m\mathbf{v} \quad \text{(L.32)}$$

leads with $U(\mathbf{v}) = e^{-i\mathbf{v}\mathbf{G}/\hbar}$ and $\mathbf{G} = t\mathbf{P} - m\mathbf{X}$, and at $t = 0$, to

$$U(\mathbf{v}) = e^{i\mathbf{v}m\mathbf{X}/\hbar}. \quad \text{(L.33)}$$

For our further considerations, we define a velocity operator $\dot{\mathbf{X}}$ by the equation $\dot{\mathbf{X}} = \frac{i}{\hbar}[H, \mathbf{X}]$. The idea is that the kinetic energy E_{kin} is determined by $E_{\text{kin}} = \frac{m\dot{\mathbf{r}}^2}{2}$. We have up to now assumed that this is the same as $\frac{\mathbf{P}^2}{2m}$, but we will soon see that this is true only under certain conditions.

Concerning the transformation behavior of the velocity operator, we assume because of (L.3) that

$$U^{-1}(\mathbf{v})\dot{\mathbf{X}}U(\mathbf{v}) \stackrel{!}{=} \dot{\mathbf{X}} + \mathbf{v}. \quad \text{(L.34)}$$

From (L.32) and (L.34), it follows that:

$$U^{-1}(\mathbf{v})\left(m\dot{\mathbf{X}} - \mathbf{P}\right)U(\mathbf{v}) = m\dot{\mathbf{X}} - \mathbf{P}. \quad \text{(L.35)}$$

This means that the operator $m\dot{\mathbf{X}} - \boldsymbol{P}$ commutes with U and therefore also with $\mathbf{X}$ (because of $U(\mathbf{v}) = e^{-i\mathbf{v}m\mathbf{X}/\hbar}$). Therefore, we can write:

$$m\dot{\mathbf{X}} - \mathbf{P} = \mathbf{f}(\mathbf{X}). \quad \text{(L.36)}$$

The relevant question is whether we can always eliminate the function $\mathbf{f}(\mathbf{X})$ by a unitary transformation. This is in fact the case in one dimension, but not in three dimensions, as we now show.

L.2.1 One-Dimensional Case

Generally, one can eliminate the function $f(X)$ by a unitary transformation only in the case of *one* dimension. Here is how: Let $F(x)$ be the anti-derivative of $f(x)$, i.e. $F'(X) = f(X)$. Then we consider the unitary transformation[41]:

$$S = e^{i\frac{F(X)}{\hbar}}. \tag{L.37}$$

Apparently, X remains unchanged under the transformation $X' = S^{-1}XS$, i.e. $X' = X$. To calculate P', we use $[P, f(X)] = -i\hbar \frac{\mathrm{d}f(X)}{dX}$ and obtain initially

$$S^{-1}PS - P = S^{-1}(PS - SP) = -i\hbar S^{-1}\frac{\mathrm{d}S(X)}{\mathrm{d}X} = -i\hbar S^{-1}\frac{i}{\hbar}f(X)S = f(X) \tag{L.38}$$

and from this, with $P' = S^{-1}PS = P + f(X)$ and (L.36), finally

$$m\dot{X}' = m\dot{X} = P + f(X) = P'. \tag{L.39}$$

In short: One can always assume that $P = m\dot{X}$; this choice is unitarily equivalent to all other choices.

With this result, we now want to determine the most general form that a Hamiltonian can take which is compatible with Galilei transformations. For this, we define the operator of the kinetic energy as

$$K = \frac{m\dot{X}^2}{2} = \frac{P^2}{2m}. \tag{L.40}$$

It follows that[42]:

$$[K, X] = \frac{1}{2m}\left[P^2, X\right] = -\frac{i\hbar}{m}P. \tag{L.41}$$

With

$$P = m\dot{X} = m\frac{i}{\hbar}[H, X], \tag{L.42}$$

we obtain

$$[H - K, X] = 0. \tag{L.43}$$

This means that $H - K$ is a function only of X, which is usually referred to as $V(X)$. Hence, the general form of a Hamiltonian which is compatible with the Galilean transformation reads in the one-dimensional case:

---

[41] Such a transformation is also called a *local gauge transformation*.

[42] Here, we again use $[x, f(p)] = i\hbar\frac{\partial f(p)}{\partial p}$.

$$H = K + V(X) = \frac{P^2}{2m} + V(X). \tag{L.44}$$

Galilei invariance manifests itself by the fact that the *form* of H is unchanged: If H is a function of X and P, then the transformed Hamiltonian H_v is the *same* function of $X' = X$ and $P' = P - mv$:

$$\begin{gathered} H = \frac{P^2}{2m} + V(X) \to \\ H' = \frac{P'^2}{2m} + V\left(X'\right) = \frac{P^2}{2m} - vP + \frac{1}{2}mv^2 + V\left(X'\right). \end{gathered} \tag{L.45}$$

L.2.2 Three-Dimensional Case

We start again with (L.36):

$$m\dot{\mathbf{X}} - \mathbf{P} = \mathbf{f}(\mathbf{X}). \tag{L.46}$$

In general, there is no transformation which causes the function $\mathbf{f}(\mathbf{x})$ to vanish. In fact, one would have to find a unitary transformation

$$S = e^{i\frac{F(\mathbf{X})}{\hbar}} \tag{L.47}$$

which satisfies

$$\mathbf{f}(\mathbf{X}) = \nabla_{\mathbf{X}} F(\mathbf{X}), \tag{L.48}$$

which is possible only for $\nabla_{\mathbf{X}} \times \mathbf{f}(\mathbf{X}) = 0$.

From (L.36), the commutation relation

$$\left[\dot{X}_i, X_j\right] = -\frac{i\hbar}{m}\delta_{ij} \tag{L.49}$$

follows. The kinetic energy is defined by

$$K = \frac{m\dot{\mathbf{X}}^2}{2} = \frac{(\mathbf{P} - \mathbf{f}(\mathbf{X}))^2}{2m}. \tag{L.50}$$

As mentioned above, we cannot assume $\mathbf{f}(\mathbf{X}) = 0$ or $m\dot{\mathbf{X}} = \boldsymbol{P}$.

For the commutator $[K, X_i]$, we have:

$$[K, X_i] = \frac{m}{2}\sum_j \left[\dot{X}_j^2, X_i\right] = \frac{m}{2}\sum_j \left(\dot{X}_j\left[\dot{X}_j, X_i\right] + \left[\dot{X}_j, X_i\right]\dot{X}_j\right) = -i\hbar\dot{X}_i. \tag{L.51}$$

Comparing this with $[H, X_i] = -i\hbar\dot{X}_i$, we can conclude that

$$[H - K, X_i] = 0. \tag{L.52}$$

This means that $H - K$ can only be a function of $\mathbf{X}$, so that we have $H = K + V(\mathbf{X})$. Thus, the most general form of a Hamiltonian which is compatible with Galilei invariance in three dimensions is:

$$H = \frac{1}{2m}\left(\mathbf{P} - \mathbf{f}\left(\mathbf{X}\right)\right)^2 + V\left(\mathbf{X}\right). \tag{L.53}$$

Note the difference between $\frac{\mathbf{P}}{m}$ and $\dot{\mathbf{X}}$; the kinetic energy reads

$$K = \frac{m\dot{\mathbf{X}}^2}{2} \neq \frac{\mathbf{P}^2}{2m}. \tag{L.54}$$

In electrodynamics, the Hamiltonian is

$$H_{cl} = \frac{1}{2m}\left(\mathbf{P} - q\mathbf{A}\right)^2 + q\varphi, \tag{L.55}$$

where $\mathbf{A}$ is the vector potential and φ the scalar potential. Thus, the elimination of $\mathbf{f}\left(\mathbf{X}\right)$ works only if $\mathbf{A}$ can be represented as a gradient, or for

$$\mathbf{B} = \nabla \times \mathbf{A} = 0, \tag{L.56}$$

i.e. for vanishing magnetic field.

L.3 Exercises

1. Given the transformation

$$U(\mathbf{v}) = e^{-i\mathbf{v}(t\mathbf{P}-m\mathbf{X})/\hbar}; \tag{L.57}$$

show that:

$$U(\mathbf{v}) = e^{-i\mathbf{v}(t\mathbf{P}-m\mathbf{X})/\hbar} = \begin{cases} e^{im\mathbf{v}\mathbf{X}/\hbar}e^{-i\mathbf{v}\mathbf{P}t/\hbar}e^{-i\frac{mv^2}{2\hbar}t} \\ e^{-i\mathbf{v}\mathbf{P}t/\hbar}e^{im\mathbf{v}\mathbf{X}/\hbar}e^{i\frac{mv^2}{2\hbar}t} \end{cases}. \tag{L.58}$$

Solution: For two operators A and B with $[A, B] = c$ (c is a complex number), it holds that $e^A e^B = e^{A+B} e^{\frac{1}{2}[A,B]}$. We consider first

$$A = \frac{im\mathbf{v}\mathbf{X}}{\hbar};\ B = -\frac{it\mathbf{v}\mathbf{P}}{\hbar}. \tag{L.59}$$

We calculate the commutator $[A, B]$ using $\left[X_i, P_j\right] = i\hbar\delta_{ij}$:

$$[A, B] = \frac{mt}{\hbar^2} ((\mathbf{vX})(\mathbf{vP}) - (\mathbf{vP})(\mathbf{vX})) = \frac{mt}{\hbar^2} \sum_{i,j} v_i v_j \left(X_i P_j - P_j X_i\right)$$
$$= \frac{mt}{\hbar^2} \sum_{i,j} v_i v_j i\hbar\delta_{ij} = i\frac{mt}{\hbar}\mathbf{v}^2. \quad \text{(L.60)}$$

With $e^{A+B} = e^{-\frac{1}{2}[A,B]}e^A e^B$, it follows that

$$U(\mathbf{v}) = e^{-i\mathbf{v}(t\mathbf{P}-m\mathbf{X})/\hbar} = e^{-i\frac{mt}{2\hbar}\mathbf{v}^2} e^{\frac{im\mathbf{vX}}{\hbar}} e^{-\frac{it\mathbf{vP}}{\hbar}}. \quad \text{(L.61)}$$

We write

$$B = \frac{im\mathbf{vX}}{\hbar};\ A = -\frac{it\mathbf{vP}}{\hbar}. \quad \text{(L.62)}$$

This leads to

$$[A, B] = -i\frac{mt}{\hbar}\mathbf{v}^2 \quad \text{(L.63)}$$

and therefore

$$U(\mathbf{v}) = e^{-i\mathbf{v}(m\mathbf{X}-t\mathbf{P})/\hbar} = e^{i\frac{mt}{2\hbar}\mathbf{v}^2} e^{-\frac{it\mathbf{vP}}{\hbar}} e^{\frac{im\mathbf{vX}}{\hbar}}. \quad \text{(L.64)}$$

2. Show the form invariance of the SEq under Galilei transformations.
Solution: If we insert $\psi(\mathbf{r}, t) = \psi\left(\mathbf{r}' - \mathbf{v}t, t\right)$ into the SEq, we obtain (L.21):

$$i\hbar\frac{\partial}{\partial t}\psi\left(\mathbf{r}' - \mathbf{v}t, t\right) + i\hbar\mathbf{v}\cdot\nabla_{\mathbf{r}'}\psi\left(\mathbf{r}' - \mathbf{v}t, t\right)$$
$$= \left[-\frac{\hbar^2}{2m}\nabla^2_{\mathbf{r}'} + V\left(\mathbf{r}' - \mathbf{v}t, t\right)\right]\psi\left(\mathbf{r}' - \mathbf{v}t, t\right). \quad \text{(L.65)}$$

We want to have $\mathbf{r}'$ instead of the argument $\mathbf{r}' - \mathbf{v}t$; in addition, we want to eliminate the term $\nabla_{\mathbf{r}'}\psi$ by a transformation. To this end, we apply the *ansatz*

$$\psi\left(\mathbf{r}' - \mathbf{v}t, t\right) = e^{iT}\varphi\left(\mathbf{r}', t\right), \quad \text{(L.66)}$$

with $T = T\left(\mathbf{r}', t\right)$ (we note that we have φ and not ψ on the right side). For the left-hand side, it follows that:

$$i\hbar\frac{\partial}{\partial t}\psi + i\hbar\mathbf{v}\cdot\nabla_{\mathbf{r}'}\psi$$
$$= i\hbar\left[e^{iT}\dot{\varphi} + i\dot{T}e^{iT}\varphi\right] + i\hbar\mathbf{v}\cdot\left[e^{iT}\nabla_{\mathbf{r}'}\varphi + i\left(\nabla_{\mathbf{r}'}T\right)e^{iT}\varphi\right]. \quad \text{(L.67)}$$

and for the Laplacian on the right-hand side,

$$\nabla_{\mathbf{r}'}^2 \psi = \nabla_{\mathbf{r}'} \left(e^{iT} \nabla_{\mathbf{r}'} \varphi \right) + i \nabla_{\mathbf{r}'} \left(\nabla_{\mathbf{r}'} T \right) e^{iT} \varphi$$
$$= e^{iT} \nabla_{\mathbf{r}'}^2 \varphi + 2i \left(\nabla_{\mathbf{r}'} T \right) e^{iT} \left(\nabla_{\mathbf{r}'} \varphi \right) + i \left(\nabla_{\mathbf{r}'}^2 T \right) e^{iT} \varphi - \left(\nabla_{\mathbf{r}'} T \right)^2 e^{iT} \varphi. \tag{L.68}$$

Inserting into (L.65) yields

$$i\hbar \left[\dot{\varphi} + i\dot{T}\varphi \right] + i\hbar \mathbf{v} \cdot \left[\nabla_{\mathbf{r}'} \varphi + i \left(\nabla_{\mathbf{r}'} T \right) \varphi \right]$$
$$= -\frac{\hbar^2}{2m} \left[\nabla_{\mathbf{r}'}^2 \varphi + 2i \left(\nabla_{\mathbf{r}'} T \right) \left(\nabla_{\mathbf{r}'} \varphi \right) + i \left(\nabla_{\mathbf{r}'}^2 T \right) \varphi - \left(\nabla_{\mathbf{r}'} T \right)^2 \varphi \right]$$
$$+ V \left(\mathbf{r}' - \mathbf{v}t, t \right) \varphi. \tag{L.69}$$

If we want to obtain as our result the SEq for φ:

$$i\hbar\dot{\varphi} = -\frac{\hbar^2}{2m} \nabla_{\mathbf{r}'}^2 \varphi + V \left(\mathbf{r}' - \mathbf{v}t, t \right) \varphi, \tag{L.70}$$

then the other terms must cancel. Sorting by coefficients of φ and $\nabla_{\mathbf{r}'} \varphi$, we initially have

$$\begin{aligned} \varphi \Rightarrow -\hbar\dot{T} - \hbar\mathbf{v} \cdot \nabla_{\mathbf{r}'} T &= -\tfrac{\hbar^2}{2m} i \left(\nabla_{\mathbf{r}'}^2 T \right) + \tfrac{\hbar^2}{2m} \left(\nabla_{\mathbf{r}'} T \right)^2 \\ \nabla_{\mathbf{r}'} \varphi \Rightarrow i\hbar\mathbf{v} &= -\tfrac{\hbar^2}{2m} 2i \nabla_{\mathbf{r}'} T. \end{aligned} \tag{L.71}$$

From the second equation, it follows that

$$\frac{\hbar}{m} \nabla_{\mathbf{r}'} T = -\mathbf{v} \text{ or } T = \frac{m}{\hbar} \left(-\mathbf{v} \cdot \mathbf{r} + F(t) \right) \tag{L.72}$$

with a still-to-be-determined function $F(t)$. Inserting T into the first equation gives

$$\dot{F} = \frac{1}{2} v^2. \tag{L.73}$$

Hence it follows that

$$T = \frac{m}{\hbar} \left(-\mathbf{v}\mathbf{r} + \frac{v^2}{2} \right) \tag{L.74}$$

and we obtain finally the relation

$$\psi \left(\mathbf{r}' - \mathbf{v}t, t \right) = e^{i \left(-m\mathbf{v}\mathbf{r}' + mv^2 t/2 \right)/\hbar} \varphi \left(\mathbf{r}', t \right). \tag{L.75}$$

The wavefunction $\varphi \left(\mathbf{r}', t \right)$, defined in such a manner in the system S', satisfies the SEq:

$$i\hbar\dot{\varphi}\left(\mathbf{r}',t\right)=-\frac{\hbar^2}{2m}\nabla_{\mathbf{r}'}^2\varphi\left(\mathbf{r}',t\right)+V\left(\mathbf{r}'-\mathbf{v}t,t\right)\varphi\left(\mathbf{r}',t\right). \tag{L.76}$$

If we finally define

$$V\left(\mathbf{r}'-\mathbf{v}t,t\right):=V'\left(\mathbf{r}',t\right), \tag{L.77}$$

then the SEq is written in the system S' as:

$$i\hbar\dot{\varphi}\left(\mathbf{r}',t\right)=-\frac{\hbar^2}{2m}\nabla_{\mathbf{r}'}^2\varphi\left(\mathbf{r}',t\right)+V'\left(\mathbf{r}',t\right)\varphi\left(\mathbf{r}',t\right). \tag{L.78}$$

3. What happens to a plane wave under a Galilei boost?
   Solution: We have in S

$$\varphi(\mathbf{r},t)=\varphi_0 e^{i(\mathbf{kr}-\omega t)}. \tag{L.79}$$

With the transformation

$$\psi'\left(\mathbf{r}',t\right)=e^{-i\frac{mv^2}{2\hbar}t}e^{im\mathbf{vr}'/\hbar}\psi\left(\mathbf{r},t\right)=e^{i\frac{mv^2}{2\hbar}t}e^{im\mathbf{vr}/\hbar}\psi\left(\mathbf{r},t\right) \tag{L.80}$$

we obtain for the transformed wave function:

$$\varphi'(\mathbf{r}',t)=\varphi_0 e^{i(\mathbf{kr}-\omega t)+i\left(m\mathbf{vr}'-mv^2t/2\right)/\hbar}. \tag{L.81}$$

We consider the exponent (without the factor i). We have:

$$\begin{aligned}\mathbf{kr}-\omega t+\frac{m}{\hbar}\mathbf{vr}'-\frac{m}{2\hbar}v^2t&=\mathbf{kr}'-\mathbf{kv}t-\frac{\hbar k^2}{2m}t+\frac{m}{\hbar}\mathbf{vr}'-\frac{m}{2\hbar}v^2t\\&=\mathbf{kr}'+\frac{m}{\hbar}\mathbf{vr}'-\frac{\hbar t}{2m}\left(k^2+\frac{2m}{\hbar}\mathbf{kv}+\frac{m^2}{\hbar^2}v^2\right)\\&=\left(\mathbf{k}+\frac{m\mathbf{v}}{\hbar}\right)\mathbf{r}'-\frac{\hbar t}{2m}\left(\mathbf{k}+\frac{m\mathbf{v}}{\hbar}\right)^2,\end{aligned} \tag{L.82}$$

and it follows that

$$\varphi'(\mathbf{r}',t)=\varphi_0 e^{i\left(\mathbf{k}'\mathbf{r}'-\omega' t\right)};\ \mathbf{k}'=\mathbf{k}+\frac{m\mathbf{v}}{\hbar},\omega'=\frac{\hbar k'^2}{2m}. \tag{L.83}$$

Appendix M
Kramers' Theorem

The theorem of Kramers deals with the consequences which time-reversal invariance has for the energy levels of a system of N quantum objects with spin $\frac{1}{2}$.

We know that for the time-reversal operator $\mathcal{T}$, it holds that $\mathcal{T}^2 = c$, with $|c| = 1$. To determine the constant c, we consider the commutator $\left[\mathcal{T}^2, \mathcal{T}\right]$. On the one hand, we have:

$$\left[\mathcal{T}^2, \mathcal{T}\right] = \mathcal{T}\left[\mathcal{T}, \mathcal{T}\right] - \left[\mathcal{T}, \mathcal{T}\right]\mathcal{T} = 0, \tag{M.1}$$

and on the other hand,

$$\left[\mathcal{T}^2, \mathcal{T}\right] = \left[c, \mathcal{T}\right] = c\mathcal{T} - \mathcal{T}c = c\mathcal{T} - c^*\mathcal{T}. \tag{M.2}$$

It follows directly that $c - c^* = 0$, and because of $|c| = 1$, we find $c = \pm 1$. One can show that for integer spin, $\mathcal{T}^2 = 1$, and for half-integer spin, $\mathcal{T}^2 = -1$.

We apply these facts to a system of N quantum objects with spin $\frac{1}{2}$. On time reversal, each spin must be considered; consequently we obtain in this case $\mathcal{T}^2 = (-1)^N$. Now let $\{|E_n; N\rangle\}$ be a CONS of this system. If we assume that H is invariant under time reversal, then also $\mathcal{T}\,|E_n; N\rangle$ must be an eigenstate of H for the energy E_n. If, on the other hand, $\mathcal{T}\,|E_n; N\rangle$ is linearly dependent on $|E_n; N\rangle$ (and this *must* be the case if there is no degeneracy), then it must hold that $\mathcal{T}\,|E_n; N\rangle = \lambda\,|E_n; N\rangle$, and it follows that:

$$\mathcal{T}^2\,|E_n; N\rangle = \mathcal{T}\lambda\,|E_n; N\rangle = \lambda^*\mathcal{T}\,|E_n; N\rangle = \lambda^*\lambda\,|E_n; N\rangle = (-1)^N\,|E_n; N\rangle\,. \tag{M.3}$$

However, this is not satisfied for odd N because of $\lambda^*\lambda > 0$. Hence, it follows that the states $|E_n; N\rangle$ are degenerate for odd N. This is the statement of the *theorem of Kramers*: If the Hamiltonian of a system of N quantum objects with spin $\frac{1}{2}$ is invariant under time reversal, then for odd N, all stationary states are degenerate (Kramers' degeneracy). It can be shown that the degree of degeneracy is even.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

Appendix N
Coulomb Energy and Exchange Energy in the Helium Atom

In this appendix, we consider in more detail the calculation of the Coulomb energy and exchange energy as addressed in Chap. 23.

The Coulomb energy is given by

$$C_{nl} = \frac{e^2}{4\pi\varepsilon_0} \int \mathrm{d}^3r_1 \mathrm{d}^3r_2 \frac{\left|\psi_{100}(\mathbf{r}_1)\,\psi_{nlm}(\mathbf{r}_2)\right|^2}{|\mathbf{r}_1 - \mathbf{r}_2|}, \tag{N.1}$$

and the exchange energy by

$$A_{nl} = \frac{e^2}{4\pi\varepsilon_0} \int \mathrm{d}^3r_1 \mathrm{d}^3r_2 \frac{\psi_{100}(\mathbf{r}_1)\,\psi_{nlm}(\mathbf{r}_2)\,\psi^*_{nlm}(\mathbf{r}_1)\,\psi^*_{100}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \tag{N.2}$$

The wavefunction reads

$$\psi_{nlm}(r) = R_{nl}(r)\, Y_l^m(\vartheta, \varphi)\,. \tag{N.3}$$

It follows that

$$C_{nl} = \frac{e^2}{4\pi\varepsilon_0} \frac{1}{4\pi} \int \mathrm{d}^3r_1 \mathrm{d}^3r_2 \frac{R^2_{10}(r_1)\,R^2_{nl}(r_2)\left|Y_l^m(\vartheta_2, \varphi_2)\right|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{N.4}$$

and

$$A_{nl} = \frac{e^2}{4\pi\varepsilon_0} \frac{1}{4\pi} \int \mathrm{d}^3r_1 \mathrm{d}^3r_2 \frac{R_{10}(r_1)\,R_{nl}(r_1)\,Y_l^{m*}(\vartheta_1, \varphi_1)\,Y_l^m(\vartheta_2, \varphi_2)\,R_{10}(r_2)\,R_{nl}(r_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \tag{N.5}$$

We have (see Appendix B, Vol. 2):

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

$$\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = \sum_{L,M} \frac{r_<^L}{r_>^{L+1}} \frac{4\pi}{2L+1} Y_L^M(\vartheta_1, \varphi_1)\, Y_L^{M*}(\vartheta_2, \varphi_2)\,. \tag{N.6}$$

Here, $r_<$ and $r_>$ denote the smaller and larger values of r_1 and r_2.

It follows that:

$$\begin{aligned} C_{nl} = {} & \frac{e^2}{4\pi\varepsilon_0} \frac{1}{4\pi} \sum_{L,M} \frac{4\pi}{2L+1} \int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 R_{10}^2(r_1)\, R_{nl}^2(r_2)\, Y_l^m(\vartheta_2, \varphi_2)\, Y_l^{m*}(\vartheta_2, \varphi_2) \\ & \times \frac{r_<^L}{r_>^{L+1}} Y_L^M(\vartheta_1, \varphi_1)\, Y_L^{M*}(\vartheta_2, \varphi_2) \end{aligned} \tag{N.7}$$

and

$$\begin{aligned} A_{nl} = {} & \frac{e^2}{4\pi\varepsilon_0} \frac{1}{4\pi} \sum_{L,M} \frac{4\pi}{2L+1} \int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 R_{10}(r_1)\, R_{nl}(r_1)\, Y_l^{m*}(\vartheta_1, \varphi_1)\, Y_l^m(\vartheta_2, \varphi_2) \\ & \times \frac{r_<^L}{r_>^{L+1}} R_{10}(r_2)\, R_{nl}(r_2)\, Y_L^M(\vartheta_1, \varphi_1)\, Y_L^{M*}(\vartheta_2, \varphi_2)\,. \end{aligned} \tag{N.8}$$

The angular part of the term C_{nl} is given by

$$\begin{aligned} C_{nl} \to {} & \sum_{L,M} \frac{4\pi}{2L+1} \int \mathrm{d}\Omega_1 \mathrm{d}\Omega_2 Y_l^m(\vartheta_2, \varphi_2)\, Y_l^{m*}(\vartheta_2, \varphi_2)\, Y_L^M(\vartheta_1, \varphi_1)\, Y_L^{M*}(\vartheta_2, \varphi_2) \\ = {} & \sum_{L,M} \frac{4\pi}{2L+1} \int Y_L^M(\vartheta_1, \varphi_1)\, \mathrm{d}\Omega_1 Y_l^m(\vartheta_2, \varphi_2)\, Y_l^{m*}(\vartheta_2, \varphi_2)\, Y_L^{M*}(\vartheta_2, \varphi_2)\, \mathrm{d}\Omega_2 \\ = {} & \sqrt{4\pi} \sum_{L,M} \frac{4\pi}{2L+1} \int Y_0^{0*}(\vartheta_1, \varphi_1)\, Y_L^M(\vartheta_1, \varphi_1)\, \mathrm{d}\Omega_1 Y_l^m(\vartheta_2, \varphi_2)\, Y_l^{m*}(\vartheta_2, \varphi_2) \\ & \times Y_L^{M*}(\vartheta_2, \varphi_2)\, \mathrm{d}\Omega_2 \\ = {} & \sum_{L,M} \frac{4\pi}{2L+1} \delta_{L0} \delta_{M0} \int Y_l^m(\vartheta_2, \varphi_2)\, Y_l^{m*}(\vartheta_2, \varphi_2)\, \mathrm{d}\Omega_2 = 4\pi \delta_{L0} \delta_{M0}, \end{aligned} \tag{N.9}$$

and for A_{nl} by

$$\begin{aligned} A_{nl} \to {} & \sum_{L,M} \frac{4\pi}{2L+1} \int \mathrm{d}\Omega_1 \mathrm{d}\Omega_2 Y_l^{m*}(\vartheta_1, \varphi_1)\, Y_l^m(\vartheta_2, \varphi_2)\, Y_L^M(\vartheta_1, \varphi_1)\, Y_L^{M*}(\vartheta_2, \varphi_2) \\ = {} & \sum_{L,M} \frac{4\pi}{2L+1} \int Y_l^{m*}(\vartheta_1, \varphi_1)\, Y_L^M(\vartheta_1, \varphi_1)\, \mathrm{d}\Omega_1 Y_l^m(\vartheta_2, \varphi_2)\, Y_L^{M*}(\vartheta_2, \varphi_2)\, \mathrm{d}\Omega_2 \\ = {} & \sum_{L,M} \frac{4\pi}{2L+1} \delta_{lL} \delta_{mM} = \frac{4\pi}{2l+1}. \end{aligned} \tag{N.10}$$

This gives

$$C_{nl} = \frac{e^2}{4\pi\varepsilon_0} \int r_1^2 \mathrm{d}r_1 r_2^2 \mathrm{d}r_2 R_{10}^2(r_1) R_{nl}^2(r_2) \frac{1}{r_>}$$
$$= \frac{e^2}{4\pi\varepsilon_0} \int r_1^2 \mathrm{d}r_1 R_{10}^2(r_1) \left[\frac{1}{r_1} \int_0^{r_1} r_2^2 R_{nl}^2(r_2)\, \mathrm{d}r_2 + \int_{r_1}^{\infty} r_2 R_{nl}^2(r_2)\, \mathrm{d}r_2 \right] \tag{N.11}$$

and

$$A_{nl} = \frac{e^2}{4\pi\varepsilon_0} \frac{1}{2l+1} \int r_1^2 \mathrm{d}r_1 r_2^2 \mathrm{d}r_2 R_{10}(r_1) R_{nl}(r_1) R_{10}(r_2) R_{nl}(r_2) \frac{r_<^l}{r_>^{l+1}}$$
$$= \frac{e^2}{4\pi\varepsilon_0} \frac{1}{2l+1} \int r_1^2 \mathrm{d}r_1 R_{10}(r_1) R_{nl}(r_1)$$
$$\times \left[\frac{1}{r_1^{l+1}} \int_0^{r_1} r_2^{2+l} R_{10}(r_2) R_{nl}(r_2)\, \mathrm{d}r_2 + r_1^l \int_{r_1}^{\infty} r_2^{1-l} R_{10}(r_2) R_{nl}(r_2)\, \mathrm{d}r_2 \right]. \tag{N.12}$$

We need the expressions for $n = 2$ and $l = 0, 1$; the radial functions (for helium, we have $Z = 2$) are given by:

$$R_{10}(r) = 2\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-\frac{Zr}{a_0}}$$
$$R_{20}(r) = 2\left(\frac{Z}{2a_0}\right)^{\frac{3}{2}} \left(1 - \frac{Zr}{2a_0}\right) e^{-\frac{Zr}{2a_0}} \tag{N.13}$$
$$R_{21}(r) = \frac{1}{\sqrt{3}}\left(\frac{Z}{2a_0}\right)^{\frac{3}{2}} \frac{Zr}{a_0} e^{-\frac{Zr}{2a_0}}.$$

After inserting the radial functions, we have to calculate the integrals in (N.11) and (N.12). One can do this by hand (essentially by partial integrations of the form $\int x^n e^{-x} = -x^n e^{-x} + n \int x^{n-1} e^{-x}$), or with the help of a computational program such as *Maple* or *Mathematica*. The result is in either case:

$$C_{20} = \frac{e^2}{4\pi\varepsilon_0} \frac{17}{81} \frac{Z}{a_0}; \quad C_{21} = \frac{e^2}{4\pi\varepsilon_0} \frac{59}{243} \frac{Z}{a_0} \tag{N.14}$$

and

$$A_{20} = \frac{e^2}{4\pi\varepsilon_0} \frac{16}{729} \frac{Z}{a_0}; \; A_{21} = \frac{e^2}{4\pi\varepsilon_0} \frac{112}{6561} \frac{Z}{a_0}. \tag{N.15}$$

Appendix O
The Scattering of Identical Particles

As might be expected, quantum scattering of identical objects[43] has certain peculiarities, which we briefly describe here.

We consider the scattering of two (identical) quantum objects in their center-of-mass system, i.e. we use center-of-mass coordinates $\mathbf{R}$ and relative coordinates $\mathbf{r}$:

$$\mathbf{R} = \frac{\mathbf{r}_1 + \mathbf{r}_2}{2}; \;\; \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2. \tag{O.1}$$

Interchanging the quantum objects results in $\mathbf{R} \to \mathbf{R}$ and $\mathbf{r} \to -\mathbf{r}$.

Now we let the two quantum objects scatter. There are two possibilities, as outlined in Fig. O.1, which we can formally describe by the transition $(r, \vartheta, \varphi) \to (r, \pi - \vartheta, \varphi + \pi)$.

We begin with the case that the two quantum objects are in principle distinguishable, although the two detectors are insensitive to this difference.[44] Then the differential scattering cross section for one quantum object to be scattered at the angle (ϑ, φ) is equal to the scattering cross section of the relative-coordinate particle into the same direction, i.e.

$$\frac{d\sigma^{(1)}}{d\Omega} = |f(\vartheta, \varphi)|^2 . \tag{O.2}$$

The scattering cross section for quantum object 2 to be scattered at the same angle is obtained by the transformation $\mathbf{r} \to -\mathbf{r}$ or $(\vartheta, \varphi) \to (\pi - \vartheta, \varphi + \pi)$:

$$\frac{d\sigma^{(2)}}{d\Omega} = |f(\pi - \vartheta, \varphi + \pi)|^2 . \tag{O.3}$$

The total differential cross section (=counting rate of the detector at an angle (ϑ, φ)) is the sum:

---

[43] Although the term 'identical particles' is familiar, we instead prefer to use the term 'quantum object' in this chapter.

[44] For example, the scattering of an electron and a muon, or of two isotopes such as ^{12}C and ^{13}C.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

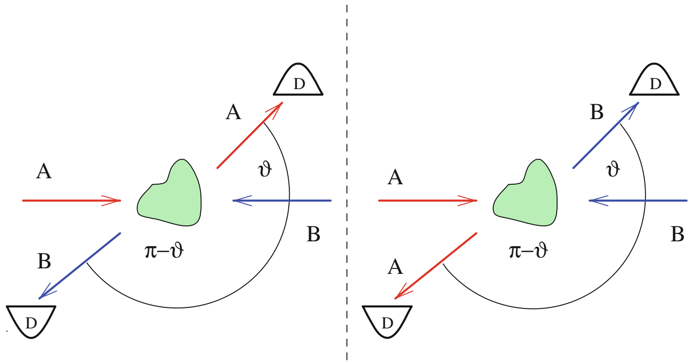


Fig. O.1 Scattering of two objects under the angles ϑ and $\pi - \vartheta$ (D = detector)

$$\frac{d\sigma}{d\Omega} = \frac{d\sigma^{(1)}}{d\Omega} + \frac{d\sigma^{(2)}}{d\Omega} = |f(\vartheta, \varphi)|^2 + |f(\pi - \vartheta, \varphi + \pi)|^2 . \tag{O.4}$$

It is crucial that we *in principle* can distinguish between the two quantum objects, regardless of whether or not we do so.

Now we suppose that the two quantum objects are identical. Then the two scattering processes in Fig. O.1 are indistinguishable even in principle. Consequently, we must perform a symmetrization of the total wavefunction (integer spin) or an antisymmetrization (half-integer spin), as described in Chap. 23.

Since the spin part is symmetric for bosons, the position part must also be symmetrized. We consider in the following two bosons with spin 0. In the case of fermions, the two quantum objects are supposed to have spin $\frac{1}{2}$. If the two fermions are in the singlet state, the spin part is antisymmetric (see Chap. 23); therefore, the spatial part must be symmetrical. In contrast, when the two fermions are in the triplet state, the spin part is symmetric and the spatial part must be antisymmetric. The bottom line is that the differential scattering cross section for a symmetrized (upper sign) or an antisymmetrized (lower sign) spatial wavefunction reads

$$\frac{d\sigma}{d\Omega} = |f(\vartheta, \varphi) \pm f(\pi - \vartheta, \varphi + \pi)|^2 . \tag{O.5}$$

In other words, we have to add the (properly symmetrized) amplitudes and not their absolute squares, as we have always done for those systems where we could not decide which path the quantum objects have taken (double slit, interaction-free quantum measurement, etc.).

For the bosons, it follows for the differential cross section (we restrict ourselves to central forces, which eliminates the dependence on φ):

$$\frac{d\sigma}{d\Omega} = |f(\vartheta) + f(\pi - \vartheta)|^2$$
$$= |f(\vartheta)|^2 + |f(\pi - \vartheta)|^2 + 2\,\mathrm{Re}\left[f(\vartheta)\, f^*(\pi - \vartheta)\right]. \tag{O.6}$$

For the two fermions with spin $\frac{1}{2}$, we distinguish between the singlet case:

$$\frac{d\sigma}{d\Omega} = |f_s(\vartheta, \varphi) + f_s(\pi - \vartheta, \varphi + \pi)|^2 ; \tag{O.7}$$

and the triplet case:

$$\frac{d\sigma}{d\Omega} = |f_t(\vartheta, \varphi) - f_t(\pi - \vartheta, \varphi + \pi)|^2 . \tag{O.8}$$

These two cases are distinguishable in principle; for an equal distribution of the states (triplet to singlet ratio $= 3/1$), the spin-insensitive detector thus measures the differential scattering cross-section

$$\frac{d\sigma}{d\Omega} = \frac{1}{4}|f_s(\vartheta, \varphi) + f_s(\pi - \vartheta, \varphi + \pi)|^2$$
$$+ \frac{3}{4}|f_t(\vartheta, \varphi) - f_t(\pi - \vartheta, \varphi + \pi)|^2 . \tag{O.9}$$

For spin-independent central forces, we find:

$$f_s(\vartheta, \varphi) = f_t(\vartheta, \varphi) = f(\vartheta) , \tag{O.10}$$

and it follows that

$$\frac{d\sigma}{d\Omega} = |f(\vartheta)|^2 + |f(\pi - \vartheta)|^2 - \mathrm{Re}\left[f(\vartheta)\, f^*(\pi - \vartheta)\right]. \tag{O.11}$$

We see the familiar distinction between distinguishable (O.4) and indistinguishable terms, namely the occurrence of interference.

Generalization of (O.6) and (O.11) for arbitrary integer and half-integer angular momenta j leads to:

$$\frac{d\sigma}{d\Omega} = |f(\vartheta)|^2 + |f(\pi - \vartheta)|^2 + 2\frac{(-1)^{2j}}{2j + 1}\mathrm{Re}\left[f(\vartheta)\, f^*(\pi - \vartheta)\right]. \tag{O.12}$$

One sees nicely how the interferences diminish with increasing angular momentum.

Appendix P
The Hadamard Transformation

We will discuss here briefly the relation between the Hadamard transformation and the Mach–Zehnder interferometer (MZI, Fig. P.1) or beam splitter (cf. Chap. 6, Vol. 1).

P.1 The MZI and the Hadamard Transformation

Mach–Zehnder interferometers are used in many different contexts, among others also in introductions to quantum information. If the details of the setup are not of interest, we can combine the effects of the two mirrors (S) and of the beam splitters (T). In this way, we can represent a MZI (i.e. TST) as a combination of two 'effective' beam splitters, i.e. as H^2 (here, H means Hadamard and *not* Hamiltonian), with

$$H^2 = TST = 1, \tag{P.1}$$

where, in the matrix representation, we have[45]:

$$T = \frac{(1+i)}{2}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix};\ S = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}. \tag{P.2}$$

To determine H, we represent the effect of the mirrors as the product of two matrices P and Q, i.e. $S = QP$, where we have to choose the matrices in such a way that $TQ = PT$ holds. In this case, we have:

$$H^2 = TST = TQPT = PTPT \text{ or } H = PT \tag{P.3}$$

[45] We omit here the distinction between an operator and its representation, i.e. between $=$ and $\cong$.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

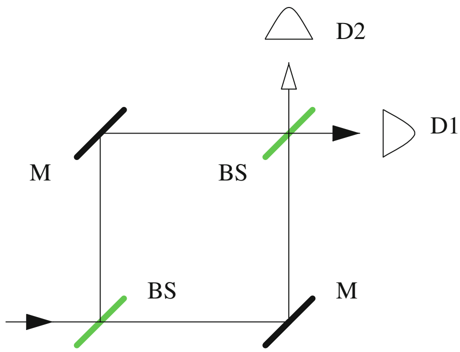


Fig. P.1 The Mach–Zehnder interferometer. M = mirror, BS = beam splitter, D = detector

With the matrices

$$P = -\frac{i}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}; \; Q = P^T = -\frac{i}{\sqrt{2}}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \tag{P.4}$$

it follows that

$$H = PT = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{P.5}$$

The unitary matrix H is called the *Hadamard matrix*[46]; with its help we can write the effect of the MZI (or the corresponding operator) as

$$H^2 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{P.6}$$

as it indeed must be.

The Hadamard matrix is related to the Pauli matrices by

$$H = \frac{\sigma_x + \sigma_z}{\sqrt{2}}. \tag{P.7}$$

P.2 The Beam Splitter and the Hadamard Transformation

To realize the Hadamard transformation *directly* by a beam splitter (BS) (i.e. without mirrors), we have to perform phase shifts. A possible setup is shown in Fig. P.2:

[46]This is a two-dimensional special case. A Hadamard matrix H_n of order n is an $n \times n$-matrix with elements $+1$ and -1, which satisfies $H_n H_n^T = nE_n$, where E_n is the n-dimensional unit matrix. A generalized Hadamard matrix can contain arbitrary elements and satisfies the equation $H_n H_n^\dagger = nE_n$.

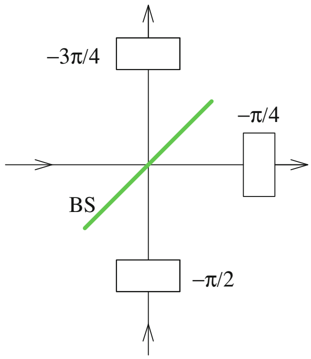


Fig. P.2 Beam splitter with phase shifters

Phase shifts of $-\pi/2$ in each case at the two vertical beams and in addition $-\pi/4$ in each case at the two outgoing beams.

The effect of the beam splitter can be written as (cf. Chap. 6, Vol. 1):

$$\begin{aligned} |H\rangle &\rightarrow \tfrac{(1+i)}{2}\left[|H\rangle + i\,|V\rangle\right] \\ |V\rangle &\rightarrow \tfrac{(1+i)}{2}\left[|V\rangle + i\,|H\rangle\right]. \end{aligned} \tag{P.8}$$

If an incoming vertical photon impinges upon a beam splitter plus phase shifter, we have (with $\frac{(1+i)}{2} = \frac{e^{i\pi/4}}{\sqrt{2}}$ and $i = e^{i\pi/2}$):

$$\begin{aligned} |V\rangle \underset{\text{phase}}{\rightarrow} e^{-i\pi/2}\,|V\rangle &\underset{\text{BS}}{\rightarrow} e^{-i\pi/2}\frac{e^{i\pi/4}}{\sqrt{2}}\left[|V\rangle + e^{i\pi/2}\,|H\rangle\right] \\ &\underset{\text{phase}}{\rightarrow} e^{-i\pi/2}\frac{e^{i\pi/4}}{\sqrt{2}}\left[e^{-3i\pi/4}\,|V\rangle + e^{i\pi/2}e^{-i\pi/4}\,|H\rangle\right], \end{aligned} \tag{P.9}$$

and for a horizontal photon:

$$|H\rangle \underset{\text{BS}}{\rightarrow} \frac{e^{i\pi/4}}{\sqrt{2}}\left[|H\rangle + e^{i\pi/2}\,|V\rangle\right] \underset{\text{phase}}{\rightarrow} \frac{e^{i\pi/4}}{\sqrt{2}}\left[e^{-i\pi/4}\,|H\rangle + e^{i\pi/2}e^{-3i\pi/4}\,|V\rangle\right]; \tag{P.10}$$

or, summarizing,

$$|H\rangle \rightarrow \frac{1}{\sqrt{2}}\left[|H\rangle + |V\rangle\right]; \quad |V\rangle \rightarrow \frac{1}{\sqrt{2}}\left[|H\rangle - |V\rangle\right] \tag{P.11}$$

as required.

P.3 The Hadamard Transformation and Quantum Information

In order to apply the notation usual in quantum information, we make the replacement $|H\rangle \rightarrow |0\rangle$, $|V\rangle \rightarrow |1\rangle$. Hence, we find with (P.5) for the Hadamard transformation:

$$|0\rangle \rightarrow \frac{|0\rangle + |1\rangle}{\sqrt{2}}; \ |1\rangle \rightarrow \frac{|0\rangle - |1\rangle}{\sqrt{2}}, \tag{P.12}$$

or, in short form:

$$|x\rangle \underset{H}{\rightarrow} \frac{|1-x\rangle + (-1)^x |x\rangle}{\sqrt{2}}; \ x = 0, 1. \tag{P.13}$$

Especially in quantum information, the Hadamard transformation is very frequently used (see Chap. 26); its specific symbol or abbreviation is $-\boxed{H}-$.

As an example of an application, we consider the preparation of special states. We assume three quantum objects, each with states $|0\rangle$ and $|1\rangle$. Initially, all three objects are in the state $|0\rangle$; thus, we have for the ground state $|0\rangle\,|0\rangle\,|0\rangle = |000\rangle$. We apply a Hadamard transformation to each individual state and obtain

$$\begin{aligned} |0\rangle\,|0\rangle\,|0\rangle &\rightarrow \frac{|0\rangle + |1\rangle}{\sqrt{2}} \frac{|0\rangle + |1\rangle}{\sqrt{2}} \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\ &= \frac{|0\rangle|0\rangle|0\rangle + |0\rangle|0\rangle|1\rangle + |0\rangle|1\rangle|0\rangle + |1\rangle|0\rangle|0\rangle + |0\rangle|1\rangle|1\rangle + |1\rangle|0\rangle|1\rangle + |1\rangle|1\rangle|0\rangle + |1\rangle|1\rangle|1\rangle}{2\sqrt{2}}. \end{aligned} \tag{P.14}$$

We see that we obtain a linear combination of *all possible* states through the application of the Hadamard transformation. In short, we can write:

$$|000\rangle \rightarrow \frac{|000\rangle + |001\rangle + |010\rangle + |100\rangle + |011\rangle + |101\rangle + |110\rangle + |111\rangle}{2\sqrt{2}}. \tag{P.15}$$

If we assume this to be the binary representation of numbers, we obtain in the decimal representation

$$\begin{aligned} |000\rangle &\underset{\text{binary}}{\rightarrow} \frac{|000\rangle + |001\rangle + |010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle + |111\rangle}{2\sqrt{2}} \\ |0\rangle &\underset{\text{decimal}}{\rightarrow} \frac{|0\rangle + |1\rangle + |2\rangle + |3\rangle + |4\rangle + |5\rangle + |6\rangle + |7\rangle}{2\sqrt{2}}. \end{aligned} \tag{P.16}$$

In general, we start from n states $|0\rangle$. If we apply a Hadamard transformation to each individual state, we can generate all the possible states of the total system. This means in decimal notation:

$$|0\rangle \underset{\text{decimal}}{\rightarrow} \sum_{m=0}^{2^n-1} |m\rangle\,. \tag{P.17}$$

If we start from an initial configuration in which the quantum objects are in the state $|0\rangle$ or $|1\rangle$, we obtain again a linear combination of all possible total states; however, the signs are different.

Appendix Q
From the Interferometer to the Computer

We want to show here that the Mach–Zehnder interferometer, introduced in Chap. 6, Vol. 1, is basically one of the essential building blocks for a quantum computer as discussed in Chap. 26. We follow in part the presentation given in arXiv:0011013.

The Mach–Zehnder Interferometer

We assume a setup which is slightly modified compared to Chap. 6, Vol. 1. First, we add a phase shifter in the upper and the lower beam paths, and secondly, we use Hadamard beam splitters (H-BS, i.e. conventional beam splitters with phase shifters added[47]); see Fig. Q.1. This setup corresponds to the overall transformation M_C:

$$M_C = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} e^{i\alpha_1} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} e^{i\alpha_0} & 0 \\ 0 & 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{Q.1}$$

or, more compactly (see the exercises):

$$M_C = -e^{i\frac{(\alpha_0+\alpha_1)}{2}}\begin{pmatrix} \cos\frac{(\alpha_0-\alpha_1)}{2} & i\sin\frac{(\alpha_0-\alpha_1)}{2} \\ -i\sin\frac{(\alpha_0-\alpha_1)}{2} & -\cos\frac{(\alpha_0-\alpha_1)}{2} \end{pmatrix}. \tag{Q.2}$$

In order to facilitate comparison with the usual notation for quantum computers, we use $|0\rangle$ and $|1\rangle$ in this appendix instead of $|H\rangle$ and $|V\rangle$. For e.g. the initial state $|0\rangle$, we can write explicitly:

$$\begin{aligned} |0\rangle \to & -e^{i\frac{(\alpha_0+\alpha_1)}{2}}\left[\cos\frac{(\alpha_0-\alpha_1)}{2}|0\rangle + i\sin\frac{(\alpha_0-\alpha_1)}{2}|1\rangle\right] \\ & = -e^{i\frac{((\alpha_0+\alpha_1)}{2})}\left[\cos\frac{\alpha}{2}|0\rangle - i\sin\frac{\alpha}{2}|1\rangle\right] ;\ \alpha = \alpha_1 - \alpha_0. \end{aligned} \tag{Q.3}$$

---

[47] The standard beam splitter (see Chap. 6, Vol. 1) corresponds to the transformation $\frac{(1+i)}{2}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$, while the Hadamard beam splitter is described by $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$; see Appendix P, Vol. 2.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

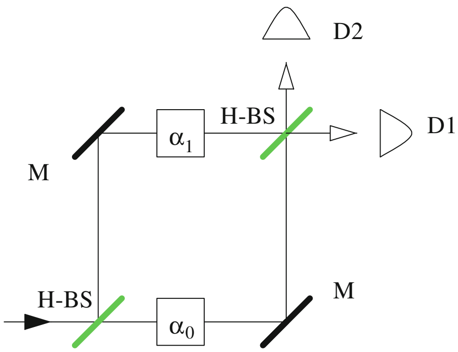


Fig. Q.1 A Mach–Zehnder interferometer, with phase shifters and Hadamard beam splitters

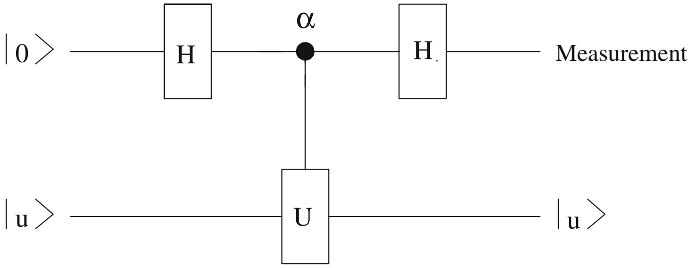
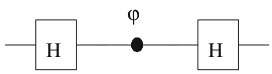


Fig. Q.2 Graphical representation of the transformation (Q.5)

Network

We restrict ourselves in the following discussion to the consideration of the input state $|0\rangle$. Then we can write the setup of Fig. Q.1 in diagram form as a network with three quantum logic gates, namely as (see the exercises):



(Q.4)

We generalize this special process of phase shifting somewhat by describing it as the application of a unitary operator U with $U\,|u\rangle = e^{i\alpha}\,|u\rangle$ in a CNOT gate as shown in Fig. Q.2. This structure is one of the basic building blocks of quantum algorithms. An input state $|0\rangle$ is transformed as follows (see the exercises):

$$|u\rangle\,|0\rangle \overset{\text{H}}{\to} |u\rangle\,\frac{|0\rangle + |1\rangle}{\sqrt{2}} \overset{\text{c-U}}{\to} |u\rangle\,\frac{|0\rangle + e^{i\alpha}\,|1\rangle}{\sqrt{2}} \overset{\text{H}}{\to} |u\rangle\,e^{i\frac{\alpha}{2}}\left[\cos\frac{\alpha}{2}\,|0\rangle - i\sin\frac{\alpha}{2}\,|1\rangle\right]. \tag{Q.5}$$

If we choose for example

$$\alpha_0 = \pm\pi; \ \alpha_1 = \alpha + \alpha_0, \tag{Q.6}$$

then this is (referred to the input state $|0\rangle$) the *exact* simulation of the Mach–Zehnder setup (Q.1), as the comparison of (Q.3) and (Q.5) shows. Note that the qubit $|u\rangle$ is not changed.

Extensions

Instead of the phase shift, we can insert any other unitary transformation for U. For example, we can choose a transformation of the controlled-U type for $f : \{0, 1\}^n \to \{0, 1\}^m$:

$$|x\rangle\,|y\rangle \to |x\rangle\,\left|[y + f(x)] \bmod 2^m\right\rangle. \tag{Q.7}$$

If we let the initial state be a superposition of all states $|y\rangle$,

$$|u\rangle = \frac{1}{2^{m/2}} \sum_{y=0}^{2^m-1} e^{-\frac{2\pi i y}{2^m}}\,|y\rangle\,, \tag{Q.8}$$

then we obtain initially

$$|x\rangle\,|u\rangle = \frac{1}{2^{m/2}} \sum_{y=0}^{2^m-1} e^{-\frac{2\pi i y}{2^m}}\,|x\rangle\,|y\rangle \to \frac{1}{2^{m/2}} \sum_{y=0}^{2^m-1} e^{-\frac{2\pi i y}{2^m}}\,|x\rangle\,\left|[y + f(x)] \bmod 2^m\right\rangle. \tag{Q.9}$$

We rearrange this expression by expanding the exponent by $f(x)$ and subsequently renaming the summation index:

$$\begin{aligned}
&\frac{1}{2^{m/2}} \sum_{y=0}^{2^m-1} e^{-\frac{2\pi i y}{2^m}}\,|x\rangle\,\left|[y + f(x)] \bmod 2^m\right\rangle \\
&= \frac{1}{2^{m/2}} e^{2\pi i \frac{f(x)}{2^m}} \sum_{y=0}^{2^m-1} e^{-2\pi i \frac{y+f(x)}{2^m}}\,|x\rangle\,\left|[y + f(x)] \bmod 2^m\right\rangle \\
&= \frac{1}{2^{m/2}} e^{2\pi i \frac{f(x)}{2^m}} \sum_{y=0}^{2^m-1} e^{-\frac{2\pi i y}{2^m}}\,|x\rangle\,|y\rangle\,,
\end{aligned} \tag{Q.10}$$

where we have used (periodicity of the complex e-function, $e^{ix} = e^{ix+2\pi i m}$):

$$\begin{aligned}
&\sum_{y=0}^{2^m-1} e^{-2\pi i \frac{y+f(x)}{2^m}}\,\left|[y + f(x)] \bmod 2^m\right\rangle \\
&\underset{z=y+f(x)}{=} \sum_{z=0}^{2^m-1} e^{-\frac{2\pi i z}{2^m}}\,|z\rangle = \sum_{y=0}^{2^m-1} e^{-\frac{2\pi i y}{2^m}}\,|y\rangle\,.
\end{aligned} \tag{Q.11}$$

Hence, (Q.10) becomes

$$|x\rangle\,|u\rangle \to e^{2\pi i \frac{f(x)}{2^m}}\,|x\rangle\,|u\rangle\,. \tag{Q.12}$$

In particular, for $m = 1$ we obtain

$$|x\rangle\,|u\rangle \to e^{\pi i f(x)}\,|x\rangle\,|u\rangle = (-1)^{f(x)}\,|x\rangle\,|u\rangle\,. \tag{Q.13}$$

Exercises

1. Determine explicitly

$$M_C = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} e^{i\alpha_1} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} e^{i\alpha_0} & 0 \\ 0 & 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{Q.14}$$

Solution: We start with

$$M_C = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & -e^{i\alpha_1} \\ -e^{i\alpha_0} & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{Q.15}$$

from which it follows that

$$M_C = -\frac{1}{2}\begin{pmatrix} e^{i\alpha_0} + e^{i\alpha_1} & e^{i\alpha_0} - e^{i\alpha_1} \\ -e^{i\alpha_0} + e^{i\alpha_1} & -e^{i\alpha_0} - e^{i\alpha_1} \end{pmatrix}. \tag{Q.16}$$

We transform with

$$M_C = -\frac{e^{i(\alpha_0+\alpha_1)/2}}{2}\begin{pmatrix} e^{i\frac{\alpha_0-\alpha_1}{2}} + e^{-i\frac{\alpha_0-\alpha_1}{2}} & e^{i\frac{\alpha_0-\alpha_1}{2}} - e^{-i\frac{\alpha_0-\alpha_1}{2}} \\ -e^{i\frac{\alpha_0-\alpha_1}{2}} + e^{-i\frac{\alpha_0-\alpha_1}{2}} & -e^{i\frac{\alpha_0-\alpha_1}{2}} - e^{-i\frac{\alpha_0-\alpha_1}{2}} \end{pmatrix} \tag{Q.17}$$

and obtain finally

$$M_C = -e^{i\frac{(\alpha_0+\alpha_1)}{2}}\begin{pmatrix} \cos\frac{(\alpha_0-\alpha_1)}{2} & i\sin\frac{(\alpha_0-\alpha_1)}{2} \\ -i\sin\frac{(\alpha_0-\alpha_1)}{2} & -\cos\frac{(\alpha_0-\alpha_1)}{2} \end{pmatrix}. \tag{Q.18}$$

2. Determine the transformation corresponding to Fig. Q.1, if standard beam splitters are used instead of the Hadamard beam splitters.
Solution: The transformation reads

$$M = \frac{(1+i)}{2}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}\begin{pmatrix} e^{i\alpha_1} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} e^{i\alpha_0} & 0 \\ 0 & 1 \end{pmatrix}\frac{(1+i)}{2}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}. \tag{Q.19}$$

It follows that:

$$M = -\frac{i}{2}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}\begin{pmatrix} 0 & e^{i\alpha_1} \\ e^{i\alpha_0} & 0 \end{pmatrix}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}, \tag{Q.20}$$

or

$$M = -\frac{i}{2}\begin{pmatrix} ie^{i\alpha_1} + ie^{i\alpha_0} & e^{i\alpha_1} - e^{i\alpha_0} \\ -e^{i\alpha_1} + e^{i\alpha_0} & ie^{i\alpha_1} + ie^{i\alpha_0} \end{pmatrix}. \tag{Q.21}$$

3. Determine the transformation corresponding to the network (Q.4) and show that it agrees for the input state $|0\rangle$ with the transformation (Q.2) up to a global phase.
Solution: The network (Q.4) corresponds to the transformation

$$\begin{aligned} &\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+e^{i\alpha} & 1-e^{i\alpha} \\ 1-e^{i\alpha} & 1+e^{i\alpha} \end{pmatrix} \\ &= \frac{1}{2}e^{i\alpha/2}\begin{pmatrix} e^{-i\alpha/2}+e^{i\alpha/2} & e^{-i\alpha/2}-e^{i\alpha/2} \\ e^{-i\alpha/2}-e^{i\alpha/2} & e^{-i\alpha/2}+e^{i\alpha/2} \end{pmatrix} = e^{i\alpha/2}\begin{pmatrix} \cos\frac{\alpha}{2} & -i\sin\frac{\alpha}{2} \\ -i\sin\frac{\alpha}{2} & \cos\frac{\alpha}{2} \end{pmatrix}. \end{aligned} \tag{Q.22}$$

Hence, the initial state $|0\rangle$ is transformed to $e^{i\alpha/2}\left[\cos\frac{\alpha}{2}|0\rangle - i\sin\frac{\alpha}{2}|1\rangle\right]$, which is identical to (Q.3) up to a global phase.
4. Derive (Q.5). How do the phases α, α_0 and α_1 have to be chosen so that the MZI of Fig. Q.1 and the network of Fig. Q.2 work identically?
Solution: The action of the Hadamard transformation is

$$|0\rangle \overset{\text{H}}{\rightarrow} \frac{|0\rangle + |1\rangle}{\sqrt{2}};\ \ |1\rangle \overset{\text{H}}{\rightarrow} \frac{|0\rangle - |1\rangle}{\sqrt{2}}, \tag{Q.23}$$

and the c-U transformation is

$$|0\rangle \overset{\text{c-U}}{\rightarrow} |0\rangle\,;\ \ |1\rangle \overset{\text{c-U}}{\rightarrow} e^{i\alpha}|1\rangle\,. \tag{Q.24}$$

Then it follows first of all that

$$|u\rangle\,|0\rangle \overset{\text{H}}{\rightarrow} |u\rangle\,\frac{|0\rangle + |1\rangle}{\sqrt{2}} \overset{\text{c-U}}{\rightarrow} |u\rangle\,\frac{|0\rangle + e^{i\alpha}|1\rangle}{\sqrt{2}}. \tag{Q.25}$$

The last Hadamard transformation acts as follows:

$$\begin{aligned} &\frac{|0\rangle + e^{i\alpha}|1\rangle}{\sqrt{2}} \overset{\text{H}}{\rightarrow} \frac{|0\rangle + |1\rangle + e^{i\alpha}\left[|0\rangle - |1\rangle\right]}{2} = \frac{\left[1+e^{i\alpha}\right]|0\rangle + \left[1-e^{i\alpha}\right]|1\rangle}{2} \\ &= \frac{e^{i\alpha/2}\left[e^{-i\alpha/2}+e^{i\alpha/2}\right]|0\rangle + e^{i\alpha/2}\left[e^{-i\alpha/2}-e^{i\alpha/2}\right]|1\rangle}{2} = e^{i\alpha/2}\left[\cos\frac{\alpha}{2}|0\rangle - i\sin\frac{\alpha}{2}|1\rangle\right]. \end{aligned} \tag{Q.26}$$

Summarized, this means that for the network,

$$|u\rangle\,|0\rangle \overset{\text{H,c-U,H}}{\rightarrow} |u\rangle\,e^{i\alpha/2}\left[\cos\frac{\alpha}{2}|0\rangle - i\sin\frac{\alpha}{2}|1\rangle\right]; \tag{Q.27}$$

and for the MZI,

$$|0\rangle \to -e^{i\frac{(\alpha_0 + \alpha_1)}{2}}\left[\cos\frac{(\alpha_1 - \alpha_0)}{2}|0\rangle - i\sin\frac{(\alpha_1 - \alpha_0)}{2}|1\rangle\right]. \tag{Q.28}$$

The square brackets are identical for $\alpha = \alpha_1 - \alpha_0$ and the phase factor for $\alpha = \alpha_0 + \alpha_1 \pm 2\pi$. Hence the two setups operate identically for

$$\alpha_0 = \pm\pi;\ \alpha_1 = \alpha + \alpha_0. \tag{Q.29}$$

Appendix R
The Grover Algorithm, Algebraically

The geometric treatment of the Grover algorithm is found in Chap. 26. To avoid back-and-forth browsing, we list here once more the essential elements from that chapter.

We assume a function $f(k)$ which vanishes for all arguments with the exception of one (the one being sought):

$$f(k) = \delta_{k\kappa};\ k = 0, 1, \ldots, N-1;\ N = 2^n;\ 0 \leq \kappa \leq N-1. \tag{R.1}$$

We use a kickback and obtain the mapping

$$|k\rangle \rightarrow (-1)^{f(k)} |k\rangle\,, \tag{R.2}$$

where $\{|k\rangle\}$ is a CONS of dimension N. Because of (R.1), this means that all states remain unchanged, except for the state being sought, $|\kappa\rangle$, where $|\kappa\rangle \rightarrow -|\kappa\rangle$ holds. Therefore, the mapping (R.2) can also be written as

$$U_\kappa = 1 - 2\,|\kappa\rangle\,\langle\kappa|\,. \tag{R.3}$$

The initial state for the algorithm is a normalized and equally-weighted superposition of all the states:

$$|s\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} |k\rangle\,;\ \ \langle s\,|s\rangle = 1. \tag{R.4}$$

With this state, we define the operator

$$U_s = 2\,|s\rangle\,\langle s| - 1. \tag{R.5}$$

Grover's algorithm is now a (repeated) application of $U_s U_\kappa$ to the initial state.

We calculate explicitly the first iteration. We apply U_κ to $|s\rangle$:

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

$$U_\kappa |s\rangle = (1 - 2|\kappa\rangle\langle\kappa|)|s\rangle = |s\rangle - 2|\kappa\rangle\langle\kappa|s\rangle = |s\rangle - \frac{2}{\sqrt{N}}|\kappa\rangle. \tag{R.6}$$

Here, we have used

$$\langle\kappa|s\rangle = \frac{1}{\sqrt{N}}. \tag{R.7}$$

In the next step, we consider U_s in addition:

$$\begin{aligned} U_s U_\kappa |s\rangle &= U_s\left(|s\rangle - \frac{2}{\sqrt{N}}|\kappa\rangle\right) = (2|s\rangle\langle s| - 1)\left(|s\rangle - \frac{2}{\sqrt{N}}|\kappa\rangle\right) \\ &= (2|s\rangle\langle s| - 1)|s\rangle - \frac{2}{\sqrt{N}}(2|s\rangle\langle s| - 1)|\kappa\rangle = \frac{N-4}{N}|s\rangle + \frac{2}{\sqrt{N}}|\kappa\rangle. \end{aligned} \tag{R.8}$$

At the outset, we had $\langle\kappa|s\rangle = 1/\sqrt{N}$; applying $U_s U_\kappa$ yields

$$\begin{aligned} \langle\kappa| U_s U_\kappa |s\rangle &= \langle\kappa|\left(\frac{N-4}{N}|s\rangle + \frac{2}{\sqrt{N}}|\kappa\rangle\right) \\ &= \frac{N-4}{N\sqrt{N}} + \frac{2}{\sqrt{N}} = \frac{3N-4}{N\sqrt{N}} = \frac{3}{\sqrt{N}}\left(1 - \frac{4}{3N}\right). \end{aligned} \tag{R.9}$$

The absolute square of the amplitude value has increased in the second step by a factor of about 9:

$$|\langle\kappa|s\rangle|^2 = \frac{1}{N};\ \ |\langle\kappa| U_s U_\kappa |s\rangle|^2 = \frac{9}{N}\left(1 - \frac{4}{3N}\right)^2 \approx \frac{9}{N}. \tag{R.10}$$

For the calculation of $\langle\kappa|(U_s U_\kappa)^m|s\rangle$ for arbitrary m, we first examine the influence of the operators on linear combinations of $|\kappa\rangle$ and $|s\rangle$. We have:

$$\begin{aligned} U_\kappa(a|\kappa\rangle + b|s\rangle) &= (1 - 2|\kappa\rangle\langle\kappa|)(a|\kappa\rangle + b|s\rangle) \\ U_s(a|\kappa\rangle + b|s\rangle) &= (2|s\rangle\langle s| - 1)(a|\kappa\rangle + b|s\rangle). \end{aligned} \tag{R.11}$$

Expanding gives:

$$\begin{aligned} U_\kappa(a|\kappa\rangle + b|s\rangle) &= a|\kappa\rangle + b|s\rangle - 2a|\kappa\rangle\langle\kappa|\kappa\rangle - 2b|\kappa\rangle\langle\kappa|s\rangle \\ U_s(a|\kappa\rangle + b|s\rangle) &= 2a|s\rangle\langle s|\kappa\rangle + 2b|s\rangle\langle s|s\rangle - a|\kappa\rangle - b|s\rangle. \end{aligned} \tag{R.12}$$

With $\langle\kappa|\kappa\rangle = 1$ and $\langle\kappa|s\rangle = 1/\sqrt{N}$, it follows that

$$\begin{aligned} U_\kappa(a|\kappa\rangle + b|s\rangle) &= a|\kappa\rangle + b|s\rangle - 2a|\kappa\rangle - 2b|\kappa\rangle\tfrac{1}{\sqrt{N}} \\ U_s(a|\kappa\rangle + b|s\rangle) &= 2a|s\rangle\tfrac{1}{\sqrt{N}} + 2b|s\rangle - a|\kappa\rangle - b|s\rangle, \end{aligned} \tag{R.13}$$

and this leads to

$$\begin{aligned} U_\kappa \left(a\left|\kappa\right\rangle + b\left|s\right\rangle\right) &= -a\left|\kappa\right\rangle + b\left|s\right\rangle - \tfrac{2b}{\sqrt{N}}\left|\kappa\right\rangle \\ U_s \left(a\left|\kappa\right\rangle + b\left|s\right\rangle\right) &= \tfrac{2a}{\sqrt{N}}\left|s\right\rangle + b\left|s\right\rangle - a\left|\kappa\right\rangle . \end{aligned} \tag{R.14}$$

Writing this in matrix form, we obtain

$$\begin{aligned} U_\kappa : \begin{pmatrix} a \\ b \end{pmatrix} &\rightarrow \begin{pmatrix} -a - \frac{2b}{\sqrt{N}} \\ b \end{pmatrix} = \begin{pmatrix} -1 & -\frac{2}{\sqrt{N}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ U_s : \begin{pmatrix} a \\ b \end{pmatrix} &\rightarrow \begin{pmatrix} -a \\ \frac{2a}{\sqrt{N}} + b \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ \frac{2}{\sqrt{N}} & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} . \end{aligned} \tag{R.15}$$

It follows[48] that

$$U_s U_\kappa = \begin{pmatrix} -1 & 0 \\ \frac{2}{\sqrt{N}} & 1 \end{pmatrix} \begin{pmatrix} -1 & -\frac{2}{\sqrt{N}} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{2}{\sqrt{N}} \\ -\frac{2}{\sqrt{N}} & 1 - \frac{4}{N} \end{pmatrix}, \tag{R.16}$$

or

$$U_s U_\kappa : \begin{pmatrix} a \\ b \end{pmatrix} \rightarrow \begin{pmatrix} a + \frac{2}{\sqrt{N}} b \\ -\frac{2}{\sqrt{N}} a + \left(1 - \frac{4}{N}\right) b \end{pmatrix}. \tag{R.17}$$

In order to calculate $(U_s U_\kappa)^m$, we first diagonalize $U_s U_\kappa$. We have (see the exercises):

$$U_s U_\kappa = \begin{pmatrix} 1 & \frac{2}{\sqrt{N}} \\ -\frac{2}{\sqrt{N}} & 1 - \frac{4}{N} \end{pmatrix} = M \begin{pmatrix} e^{2i\varphi} & 0 \\ 0 & e^{-2i\varphi} \end{pmatrix} M^{-1} \tag{R.18}$$

with

$$\varphi = \arcsin \frac{1}{\sqrt{N}} \tag{R.19}$$

and

$$M = \begin{pmatrix} -i & i \\ e^{i\varphi} & e^{-i\varphi} \end{pmatrix};\ M^{-1} = \frac{1}{2i \cos \varphi} \begin{pmatrix} -e^{-i\varphi} & i \\ e^{i\varphi} & i \end{pmatrix}. \tag{R.20}$$

It follows that:

$$(U_s U_\kappa)^m = M \begin{pmatrix} e^{2im\varphi} & 0 \\ 0 & e^{-2im\varphi} \end{pmatrix} M^{-1} \tag{R.21}$$

A small calculation (see the exercises) shows

$$(U_s U_\kappa)^m = \frac{1}{\cos \varphi} \begin{pmatrix} \cos (2m - 1)\varphi & \sin 2m\varphi \\ -\sin 2m\varphi & \cos (2m + 1)\varphi \end{pmatrix}, \tag{R.22}$$

[48]For simplicity, we do not distinguish in this section between an operator and its matrix representation, i.e. we always write $=$ instead of $\cong$.

or

$$(U_s U_\kappa)^m : \begin{pmatrix} a \\ b \end{pmatrix} \to \frac{1}{\cos\varphi} \begin{pmatrix} a\cos(2m-1)\varphi + b\sin 2m\varphi \\ -a\sin 2m\varphi + b\cos(2m+1)\varphi \end{pmatrix}, \qquad \text{(R.23)}$$

or

$$(U_s U_\kappa)^m (a|\kappa\rangle + b|s\rangle) = \frac{1}{\cos\varphi}\big([a\cos(2m-1)\varphi + b\sin 2m\varphi]\,|\kappa\rangle + [-a\sin 2m\varphi + b\cos(2m+1)\varphi]\,|s\rangle \big). \qquad \text{(R.24)}$$

We then find from this:

$$(U_s U_\kappa)^m |s\rangle = \frac{1}{\cos\varphi}\left(\sin 2m\varphi \cdot |\kappa\rangle + \cos(2m+1)\varphi \cdot |s\rangle\right), \qquad \text{(R.25)}$$

and therefore (see the exercises),

$$\begin{aligned} \langle\kappa|\,(U_s U_\kappa)^m |s\rangle &= \frac{1}{\cos\varphi}\,\langle\kappa|\left(\sin 2m\varphi \cdot |\kappa\rangle + \cos(2m+1)\varphi \cdot |s\rangle\right) \\ &= \frac{1}{\cos\varphi}\left(\sin 2m\varphi + \frac{1}{\sqrt{N}}\cos(2m+1)\varphi\right) = \sin(2m+1)\varphi. \end{aligned} \qquad \text{(R.26)}$$

As a test, we check this result for $m = 1$. We find:

$$\begin{aligned} \langle\kappa|\,(U_s U_\kappa)\,|s\rangle &= \sin 3\varphi = 3\sin\varphi - 4\sin^3\varphi \\ &= 3\frac{1}{\sqrt{N}} - 4\frac{1}{N\sqrt{N}} = \frac{3N-4}{N\sqrt{N}} \end{aligned} \qquad \text{(R.27)}$$

as expected; see (R.9).

Exercises

1. Show that:

$$\begin{pmatrix} 1 & \frac{2}{\sqrt{N}} \\ -\frac{2}{\sqrt{N}} & 1-\frac{4}{N} \end{pmatrix} = M \begin{pmatrix} e^{2i\varphi} & 0 \\ 0 & e^{-2i\varphi} \end{pmatrix} M^{-1} \qquad \text{(R.28)}$$

with $\varphi = \arcsin\frac{1}{\sqrt{N}}$ and

$$M = \begin{pmatrix} -i & i \\ e^{i\varphi} & e^{-i\varphi} \end{pmatrix};\; M^{-1} = \frac{1}{ie^{-i\varphi} + ie^{i\varphi}} \begin{pmatrix} -e^{-i\varphi} & i \\ e^{i\varphi} & i \end{pmatrix}. \qquad \text{(R.29)}$$

Solution: We have

$$
\begin{aligned}
M\begin{pmatrix} e^{2i\varphi} & 0 \\ 0 & e^{-2i\varphi} \end{pmatrix} M^{-1} &= \frac{1}{ie^{-i\varphi}+ie^{i\varphi}} \begin{pmatrix} -i & i \\ e^{i\varphi} & e^{-i\varphi} \end{pmatrix} \begin{pmatrix} e^{2i\varphi} & 0 \\ 0 & e^{-2i\varphi} \end{pmatrix} \begin{pmatrix} -e^{-i\varphi} & i \\ e^{i\varphi} & i \end{pmatrix} \\
&= \frac{1}{ie^{-i\varphi}+ie^{i\varphi}} \begin{pmatrix} ie^{i\varphi}+ie^{-i\varphi} & e^{2i\varphi}-e^{-2i\varphi} \\ -e^{2i\varphi}+e^{-2i\varphi} & ie^{3i\varphi}+ie^{-3i\varphi} \end{pmatrix} \\
&= \begin{pmatrix} 1 & \frac{1}{i}\left(e^{i\varphi}-e^{-i\varphi}\right) \\ -\frac{1}{i}\left(e^{i\varphi}-e^{-i\varphi}\right) & e^{2i\varphi}+e^{-2i\varphi}-1 \end{pmatrix}.
\end{aligned} \tag{R.30}
$$

With

$$
\begin{aligned}
e^{i\varphi} &= \cos \arcsin \frac{1}{\sqrt{N}} + i \sin \arcsin \frac{1}{\sqrt{N}} = \sqrt{1-\frac{1}{N}} + i\frac{1}{\sqrt{N}}; \\
e^{2i\varphi} &= 1-\frac{2}{N} + 2i\frac{1}{\sqrt{N}}\sqrt{1-\frac{1}{N}},
\end{aligned} \tag{R.31}
$$

it follows finally that

$$
M\begin{pmatrix} e^{2i\varphi} & 0 \\ 0 & e^{-2i\varphi} \end{pmatrix} M^{-1} = \begin{pmatrix} 1 & \frac{1}{i}2i\frac{1}{\sqrt{N}} \\ -\frac{1}{i}2i\frac{1}{\sqrt{N}} & 2-\frac{4}{N}-1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{2}{\sqrt{N}} \\ -\frac{2}{\sqrt{N}} & 1-\frac{4}{N} \end{pmatrix}. \tag{R.32}
$$

2. Show that

$$
(U_s U_\kappa)^m = \frac{1}{\cos\varphi} \begin{pmatrix} \cos(2m-1)\varphi & \sin 2m\varphi \\ -\sin 2m\varphi & \cos(2m+1)\varphi \end{pmatrix}. \tag{R.33}
$$

Solution: We start from

$$
(U_s U_\kappa)^m = M \begin{pmatrix} e^{2im\varphi} & 0 \\ 0 & e^{-2im\varphi} \end{pmatrix} M^{-1} \tag{R.34}
$$

with

$$
M = \begin{pmatrix} -i & i \\ e^{i\varphi} & e^{-i\varphi} \end{pmatrix};\; M^{-1} = \frac{1}{2i\cos\varphi} \begin{pmatrix} -e^{-i\varphi} & i \\ e^{i\varphi} & i \end{pmatrix}. \tag{R.35}
$$

It follows that:

$$
\begin{aligned}
(U_s U_\kappa)^m &= \frac{1}{2i\cos\varphi} \begin{pmatrix} -i & i \\ e^{i\varphi} & e^{-i\varphi} \end{pmatrix} \begin{pmatrix} e^{2im\varphi} & 0 \\ 0 & e^{-2im\varphi} \end{pmatrix} \begin{pmatrix} -e^{-i\varphi} & i \\ e^{i\varphi} & i \end{pmatrix} \\
&= \frac{1}{2i\cos\varphi} \begin{pmatrix} ie^{i(2m-1)\varphi}+ie^{-i(2m-1)\varphi} & e^{2im\varphi}-e^{-2im\varphi} \\ -e^{2im\varphi}+e^{-2im\varphi} & ie^{i(2m+1)\varphi}+ie^{-i(2m+1)\varphi} \end{pmatrix} \\
&= \frac{1}{\cos\varphi} \begin{pmatrix} \cos(2m-1)\varphi & \sin 2m\varphi \\ -\sin 2m\varphi & \cos(2m+1)\varphi \end{pmatrix}.
\end{aligned} \tag{R.36}
$$

3. Given that

$$\langle\kappa| \left(U_s U_\kappa\right)^m |s\rangle = \frac{1}{\cos\varphi}\left(\sin 2m\varphi + \frac{1}{\sqrt{N}}\cos\left(2m + 1\right)\varphi\right); \qquad \text{(R.37)}$$

show that

$$\langle\kappa| \left(U_s U_\kappa\right)^m |s\rangle = \sin\left(2m + 1\right)\varphi \qquad \text{(R.38)}$$

holds.
Solution: With $\sin\varphi = \frac{1}{\sqrt{N}}$ and the relevant theorems for the trigonometric functions, we find:

$$\begin{aligned} &\frac{1}{\cos\varphi}\left(\sin 2m\varphi + \frac{1}{\sqrt{N}}\cos\left(2m + 1\right)\varphi\right) \\ &\quad = \frac{1}{\cos\varphi}\left(\sin 2m\varphi + \sin\varphi\cos\left(2m + 1\right)\varphi\right) \\ &\quad = \frac{1}{\cos\varphi}\frac{\sin\left(2m\varphi\right) + \sin\left(\left(2m + 2\right)\varphi\right)}{2} \\ &\quad = \frac{1}{\cos\varphi}\sin\frac{2m\varphi + \left(2m + 2\right)\varphi}{2}\cdot\cos\frac{2m\varphi - \left(2m + 2\right)\varphi}{2} \\ &\quad = \frac{1}{\cos\varphi}\sin\left(2m + 1\right)\varphi\cdot\cos\varphi = \sin\left(2m + 1\right)\varphi. \qquad \text{(R.39)} \end{aligned}$$

Appendix S
Shor Algorithm

Shor's algorithm serves to decompose very large numbers into their prime factors. We first discuss the classical and then the quantum-mechanical part of the algorithm.

S.1 Classical Part

Given a number N (odd, not prime) whose prime factorization is to be determined, $N = p_1^{\alpha_1} p_2^{\alpha 2} \ldots p_m^{\alpha_m}$. To accomplish this, we choose at random a number a with $2 \leq a \leq N - 1$, which is relatively prime to N (otherwise we would have found a divisor of N), i.e. $gcd\,(a, N) = 1$.[49] We consider

$$a^j \bmod N; \;\; j = 0, 1, 2, \ldots \tag{S.1}$$

Beginning with a value j_p of j which depends on a and N, $a^j \bmod N$ is periodic with the *order* r (also called the *period*)[50]:

$$a^{j+r} \bmod N = a^j \bmod N \tag{S.2}$$

where r is the smallest number for which this equation is satisfied. Table S.1 shows some examples.

As can be seen, the period can be even or odd; j_p can be 0, but also assumes other values.

With the help of the period r we can determine factors of N. This works as follows: We assume that we have determined the period r (actually, this is the business of quantum mechanics) and that it meets the following conditions:

[49] gcd is an acronym for 'greatest common divisor'. There are very effective methods for determining the gcd.

[50] Some remarks on modular arithmetic are to be found below.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

Table S.1 Some examples for $a^j \bmod N$

$j =$	0	1	2	3	4	5	6	...	r	j_p
$a = 8, N = 21$	1	8	1	8	1	8	1	...	2	0
$a = 13, N = 35$	1	13	29	27	1	13	29	...	4	0
$a = 19, N = 35$	1	19	11	34	16	24	1	...	6	0
$a = 4, N = 63$	1	4	16	1	4	16	1	...	3	0
$a = 6, N = 63$	1	6	36	27	36	27	36	...	2	2
$a = 7, N = 63$	1	7	49	28	7	49	28	...	3	1

$$(1)\ r \text{ is even; } (2)\ a^{\frac{r}{2}} \bmod N \neq 1. \tag{S.3}$$

As one can show, the inequality $w \geq 1 - \frac{1}{2^{(m-1)}}$ holds for the probability w that these two conditions are satisfied for a number $N = p_1^{\alpha_1} p_2^{\alpha 2} \ldots p_m^{\alpha_m}$. If the conditions (S.3) are not met, we select a different a and start the process again. Numerical examples are given in the exercises.

Under the assumptions of (S.3), one can show that the two expressions

$$d_{\pm} = gcd\left(a^{\frac{r}{2}} \pm 1, N\right) \tag{S.4}$$

yield factors of N. One divides N by these factors and applies, if necessary, the same procedures at $\frac{N}{d_- d_+}$ to identify any other factors.

An example: Let $N = 21$; we choose $a = 8$. Then we have $r = 2$ (see the above table). With the two expressions

$$d_- = gcd\,(7, 21) = 7;\ \ d_+ = gcd\,(9, 21) = 3 \tag{S.5}$$

we have found the two factors of $21 = 3 \cdot 7$. Other examples are given below and in the exercises.

S.2 Quantum-Mechanical Part

The quantum-mechanical part is confined to the determination of the period for given N and a. We set up two registers, the first register (argument register) of length m, and the second (function register) of length L. Usually one chooses $N^2 \leq M = 2^m < 2N^2$ and $L \gtrsim \log_2 N$.

In the following, we illustrate each step with the specific example $N = 35, a = 13$ and[51] $r = 4$. For this example, $j_p = 0$, and we restrict the general formalism to this case (which implies $a^r \bmod N = 1$). The extension to $j_p > 0$ is simple, but it leads to

[51] In a 'real' problem, of course, one does not know the period; it is a number being sought.

more paperwork. Similarly, in the examples explicitly worked out, we choose for M not the just as the 'usual' value denoted by $M = 2048 = 2^{11}$, but confine ourselves for clarity[52] to $M = 128 = 2^7$.

First step: We prepare the state $|0\rangle \otimes |0\rangle \equiv |0\rangle\,|0\rangle$ (i.e. argument register $\otimes$ function register) and then apply the Hadamard transformation $H_2^{\otimes m}$ to the argument register; this yields the superposition (see Chap. 26):

$$|\varphi_1\rangle = \frac{1}{\sqrt{M}} \sum_{j=0}^{M-1} |j\rangle\,|0\rangle\,. \tag{S.6}$$

This state (a product state) contains all $M = 2^m$ numbers from 0 to $M - 1$ at the same time (it must of course be guaranteed that $M > r$ holds).

In our example with $N = 35$, $a = 13$ and $M = 128$, this means that:

$$|\varphi_1\rangle = \frac{1}{\sqrt{128}} \left(|0\rangle + |1\rangle + \cdots |127\rangle\right) |0\rangle\,. \tag{S.7}$$

Second step: We modify the function register by the unitary transformation[53] $|j\rangle\,|0\rangle \to |j\rangle\,\left|a^j \bmod N\right\rangle$. Then we obtain the total state (it is obviously an entangled state):

$$|\varphi_2\rangle = \frac{1}{\sqrt{M}} \sum_{j=0}^{M-1} |j\rangle\,\left|a^j \bmod N\right\rangle. \tag{S.8}$$

In this way, the quantum computer has calculated *in one passage* the possible values of $a^j \bmod N$ (quantum parallelism). This is the step that gives significant time savings compared to traditional methods.

We still have to read out the relevant information, i.e. the period r, but the essential part of the computational work is done. In order to see this more clearly, we re-sort. We know that $a^j \bmod N$ is periodic with period r; we can therefore write

$$j = J + kr;\ J = 0, 1, 2, \ldots, k = 0, 1, 2, \ldots \tag{S.9}$$

The numbers J are called the offset. In our example ($N = 35$, $a = 13$, $r = 4$)), the offset can take on the four values $J = 0, 1, 2, 3$.

Thus we can write:

$$a^j \bmod N = a^{J+kr} \bmod N = a^J \bmod N \ \text{ independently of } k, \tag{S.10}$$

[52] If one applies the Shor algorithm to numbers that are suitable for encryption purposes, one encounters of course some more pitfalls than are seen here. But since we just want to outline the principle of the method here, toy examples will do nicely.

[53] The exact form of the transformation is not important. It is enough to know that we can construct it by suitable combinations of H, Φ and C (Hadamard, phase, CNOT), and that it is reversible.

and it follows that

$$|\varphi_2\rangle = \frac{1}{\sqrt{M}} \sum_J \sum_{k=0}^{s-1} |J + kr\rangle \left|a^J \bmod N\right\rangle \text{ with } s = 1 + \left[\frac{M-1-J}{r}\right], \quad \text{(S.11)}$$

where $\left[\frac{p}{q}\right]$ denotes the integer portion of $\frac{p}{q}$ (also called floor or integer part). The outer sum runs over all possible offsets[54] J. Thus, we have a *superposition* of states of the argument register as 'prefactor' of a possible value of the function register.

In our example, $r = 4$, and we obtain for (S.8) the expression:

$$\begin{aligned} |\varphi_2\rangle &= \frac{1}{\sqrt{128}} \sum_{j=0}^{127} |j\rangle \left|13^j \bmod 35\right\rangle \\ &= \frac{|0\rangle|1\rangle + |1\rangle|13\rangle + |2\rangle|29\rangle + |3\rangle|27\rangle + |4\rangle|1\rangle + \cdots}{\sqrt{128}}. \end{aligned} \quad \text{(S.12)}$$

Re-sorting yields the formulations corresponding to (S.11):

$$\begin{aligned} |\varphi_2\rangle = &\frac{|0\rangle + |4\rangle + |8\rangle + \cdots + |124\rangle}{\sqrt{128}} |1\rangle + \frac{|1\rangle + |5\rangle + |9\rangle + \cdots + |125\rangle}{\sqrt{128}} |13\rangle \\ &+ \frac{|2\rangle + |6\rangle + |10\rangle + \cdots + |126\rangle}{\sqrt{128}} |29\rangle + \frac{|3\rangle + |7\rangle + |11\rangle + \cdots + |127\rangle}{\sqrt{128}} |27\rangle . \end{aligned} \quad \text{(S.13)}$$

With $M = 128$ and $r = 4$, it follows that $s = \left[\frac{128}{r}\right] = 32$. The crucial point here is that *all* superpositions are 4-periodic—only their offsets are different.

At this point, one could determine the period by measuring often enough the argument register and the function register, and sorting the results according to the different measured values. However, our toy example with $r = 4$ is insofar somewhat misleading in that it is very transparent due to the small numbers used. In 'real' problems, one is dealing with very large numbers, and accordingly the periods can be very large.[55] It would be an important aid if each function register had some superposition of the *same* states (apart from phases); in (S.11), the function registers differ due to the different offsets. This crucial simplification is achieved with the quantum Fourier transform (QFT, see Appendix H, Vol. 1, 'Fourier transforms and the delta function').

[54] The length L of the function register must be greater than or equal to the number of different offsets; therefore, we have the condition $L \gtrsim \log_2 N$.

[55] Even toy examples with 'small' numbers can lead to remarkable period lengths. As an example, we consider the prime factorization of $N = 2149841$. For $a = 3$, we obtain the period $r = 213330$. With $gcd\left(3^{106665} - 1, 2149841\right) = 131$ and $gcd\left(3^{106665} + 1, 2149841\right) = 16411$, the factorization $N = 2149841 = 131 \cdot 16411$ follows.

Third step: We apply the QFT (acting on the argument register):

$$U_{QFT} = \frac{1}{\sqrt{M}} \sum_{n,l=0}^{M-1} e^{2\pi i \frac{nl}{M}} \left|l\right\rangle \left\langle n\right| \tag{S.14}$$

to $|\varphi_2\rangle$ in (S.11) and obtain (due to $\langle n|\, j\rangle = \delta_{jn}$) the relationship:

$$\begin{aligned}
|\varphi_3\rangle &= U_{QFT}\, |\varphi_2\rangle \\
&= \frac{1}{\sqrt{M}} \sum_{n,l=0}^{M-1} e^{2\pi i \frac{nl}{M}} \left|l\right\rangle \left\langle n\right| \frac{1}{\sqrt{M}} \sum_{J} \sum_{k=0}^{s-1} |J + kr\rangle \left|a^J \bmod N\right\rangle \\
&= \frac{1}{M} \sum_{J} \sum_{k=0}^{s-1} \sum_{n,l=0}^{M-1} e^{2\pi i \frac{nl}{M}} \left|l\right\rangle \delta_{n,J+kr} \left|a^J \bmod N\right\rangle \\
&= \frac{1}{M} \sum_{J} \sum_{l=0}^{M-1} \sum_{k=0}^{s-1} e^{2\pi i \frac{(J+kr)l}{M}} \left|l\right\rangle \left|a^J \bmod N\right\rangle .
\end{aligned} \tag{S.15}$$

We have thus achieved our goal that the same states always occur in the argument registers, independently of the offset; in the expression (S.11) for $|\varphi_2\rangle$, this was still not the case. In order to formulate this more clearly, we use the fact that $\left|a^J \bmod N\right\rangle$ in (S.15) does not depend on k, and we thus can carry out the sum over k. With (see the exercises for the calculation):

$$c\,(J, r, s; l) := \sum_{k=0}^{s-1} e^{2\pi i \frac{(J+kr)l}{M}} = \begin{cases} e^{2\pi i \frac{Jl}{M}} e^{\pi i \frac{rl(s-1)}{M}} \dfrac{\sin \pi \frac{rl}{M} s}{\sin \pi \frac{rl}{M}} \,; & \frac{rl}{M} \notin \mathbb{N}_0 \\ e^{2\pi i \frac{Jl}{M}} \cdot s & \frac{rl}{M} \in \mathbb{N}_0 \end{cases} \tag{S.16}$$

we can write (S.15) as

$$|\varphi_3\rangle = \frac{1}{M} \sum_{J} \sum_{l=0}^{M-1} c\,(J, r, s; l) \left|l\right\rangle \left|a^J \bmod N\right\rangle ; \;\; s = 1 + \left[\frac{M-1-J}{r}\right]. \tag{S.17}$$

The crucial point is that $|c\,(J, r, s; l)|$, according to (S.16), is not only *independent* of J (i.e. of the function register), but also has *distinct maxima* for particular values of l, namely for l values which are multiples of $\frac{M}{r}$ (or at least approximate multiples). In other words, the QFT (S.14) causes an *amplitude amplification*, which leads to the result that measurements yield mainly l values with $l \approx n \cdot \frac{M}{r}$ with a high probability.

We illustrate these findings by means of the example $M = 128$, $r = 4$. It follows first of all that

$$s = 1 + \left[\frac{128-1-J}{4}\right] = 1 + 31 + \left[\frac{3-J}{4}\right] = 32 \;\text{ due to }\; J = 0, 1, 2, 3, \tag{S.18}$$

and from this (note: $0 \le l \le M - 1$) with (S.16):

$$|c(J,r,s;l)| = \left| \frac{\sin \pi l}{\sin \frac{\pi l}{32}} \right| = \begin{cases} 32 & \text{for } l = 0, 32, 64, 96 \\ 0 & \text{otherwise.} \end{cases} \tag{S.19}$$

Accordingly, we see in the argument register only states $|n \cdot 32\rangle$; all others vanish exactly. In this transparent example, we see immediately that the period is $r = 4$; we can formally determine it by $\frac{rl}{M} \in \mathbb{N}_0$ or $r = \frac{M}{l} n_l$. Since 32 is the greatest common divisor of (32, 64, 96), the period is given by $\frac{128}{32} = 4$.

A similar behavior always results if the period r is a power of 2 (because of $M = 2^m$, M is also a power of two). Only the states $\left|n \cdot \frac{M}{r}\right\rangle$ with $n \in \mathbb{N}_0$ survive, all others vanish exactly. If r is not a power of two, the 'unwanted' states do not disappear exactly, but their amplitudes are much smaller than those of the desired states. As an example, we choose $M = 128$ and $r = 6$. In this case, we find:

$$s = 1 + \left[\frac{128 - 1 - J}{6}\right] = 22 + \left[\frac{1 - J}{6}\right] = \begin{cases} 22 \text{ for } J = 0, 1 \\ 21 \text{ for } J = 2, 3, 4, 5 \end{cases} \tag{S.20}$$

and thus

$$|c(J,r,s;l)| = \begin{cases} \left| \frac{\sin \pi \frac{3l}{64} s}{\sin \pi \frac{3l}{64}} \right|; & \frac{3l}{64} \notin \mathbb{N}_0 \\ s & \frac{3l}{64} \in \mathbb{N}_0 \end{cases}; \ l = 0, 1, 2, \ldots, 127. \tag{S.21}$$

Obviously, $\frac{3l}{64}$ is an integer only for $l = 0$ and $l = 64$, but this holds true approximately also for the values $l = 21$ and 22, 42 and 43, 85 and 86, 106 and 107. This manifests itself in the fact that the corresponding coefficients $|c(J,r,s;l)|$ are relatively large; see Fig. S.1. Accordingly, when measuring one obtains one of these states with a high probability. Thus, we have here another non-trivial example of amplitude amplification, as discussed in Chap. 26. From Fig. S.1, the period 6 can be read out directly; formally, we can calculate it as $r \approx n \cdot \frac{M}{l}$. A more detailed analysis of this example is found in the exercises.

We summarize: *Before* the QFT, the states of the argument register depend on the function register, as is seen in the example of (S.13). *After* the QFT, the argument registers contain superpositions of the *same* states, independently of the function register; cf. (S.15). In addition, the QFT causes the measurement probabilities for states in the argument registers to become unequally distributed, as in (S.13), but with pronounced maxima at $l \approx n\frac{M}{r}$ with $n = 0, 1, 2, \ldots$. In other words, with a high probability (with certainty for $r = 2^N$), *each* measurement yields one of those states in the argument register from which the period can be determined directly. In this manner, the quantum computer determines the periodicity of a function in a few steps.

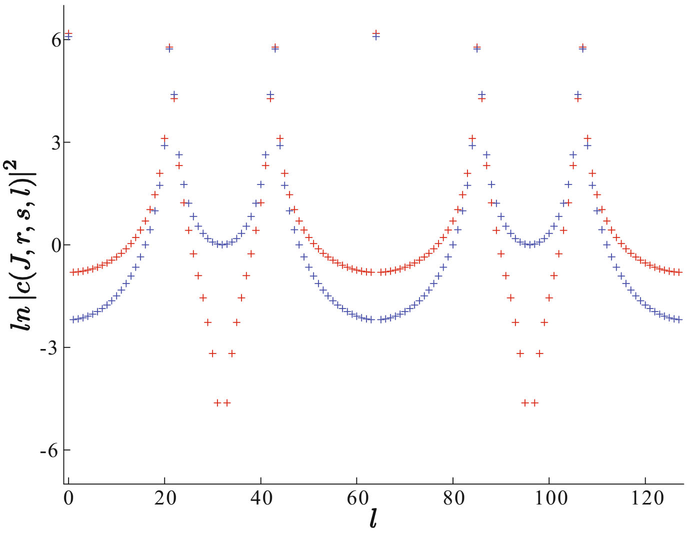


Fig. S.1 The coefficient (S.21) for $M = 128$ and $r = 6$. *Red crosses* for $s = 22$, *blue crosses* for $s = 21$

As we said above, this works with very high probability. Should it not work in a particular case (or if the conditions (S.3) are not satisfied), one chooses another number a and lets the algorithm run through once more.

S.3 Supplement on Modular Arithmetic

We present here some information on modular arithmetic. We have

$$a = p \cdot n + r \rightarrow a \bmod n = r \;\text{ or }\; a \bmod n = a - \left[\frac{a}{n}\right] n, \tag{S.22}$$

where all numbers are $\in \mathbb{N}_0$; $\left[\frac{a}{n}\right]$ is the integer part. Example: $a = 51, n = 7$.

$$\begin{aligned} 51 \bmod 7 &= (7 \cdot 7 + 2) \bmod 7 = 2 \\ 51 \bmod 7 &= 51 - \left[\frac{51}{7}\right] \cdot 7 = 51 - 7 \cdot 7 = 2. \end{aligned} \tag{S.23}$$

Modular addition and multiplication are defined by

$$(a + b) \bmod n = (a \bmod n + b \bmod n) \bmod n \tag{S.24}$$

and

$$ab \bmod n = [(a \bmod n) \cdot (b \bmod n)] \bmod n. \tag{S.25}$$

Example:

$$\begin{aligned}(52 + 34) \bmod 7 = 86 \bmod 7 = 2 &= (52 \bmod 7 + 34 \bmod 7) \bmod 7 \\ &= (3 + 6) \bmod 7 = 2\end{aligned} \tag{S.26}$$

and

$$\begin{aligned}(52 \cdot 34) \bmod 7 = 1768 \bmod 7 = 4 &= (52 \bmod 7 \cdot 34 \bmod 7) \bmod 7 \\ &= (3 \cdot 6) \bmod 7 = 4.\end{aligned} \tag{S.27}$$

If the calculations refer to only *one* n, the following shorthand notation is often used in quantum information applications:

$$a \oplus b := (a + b) \bmod n \tag{S.28}$$

where here the symbol $\oplus$ of course does *not* denote the direct sum of vectors.

Because of (S.25), powers can be calculated recursively:

$$a^{r+1} \bmod N = \left[\left(a^r \bmod N\right) \cdot (a \bmod N)\right] \bmod N. \tag{S.29}$$

S.4 Exercises

1. Given two integers n and a, calculate the period $r = a \bmod n$, and with (S.4) the factors of n.

   (a) $n = 35$ and $a = 13$
   Solution: We have

13^0	$= 1$	$= 0 \cdot 35 + 1$	$= 1 \pmod{35}$
13^1	$= 13$	$= 0 \cdot 35 + 13$	$= 13 \pmod{35}$
13^2	$= 169$	$= 4 \cdot 35 + 29$	$= 29 \pmod{35}$
13^3	$= 2197$	$= 62 \cdot 35 + 27$	$= 27 \pmod{35}$
13^4	$= 28561$	$= 816 \cdot 35 + 1$	$= 1 \pmod{35}$
13^5	$= 371293$	$= 10608 \cdot 35 + 13$	$= 13 \pmod{35}$

Hence, the period is $r = 4$. It follows that $a^{r/2} = 13^2 = 169$ and thus

$$gcd\,(168, 35) = 7;\;\; gcd\,(170, 35) = 5;\;\; 7 \cdot 5 = 35 \tag{S.30}$$

Table S.2 Solution of exercise 2 (a), $N = 21$

a	r	$a^{r/2}$	$gcd(a^{r/2} - 1, 15)$	$gcd(a^{r/2} + 1, 15)$
2	6	$2^3 = 8$	$(7, 21) = 7$	$(9, 21) = 3$
4	3			
5	6	$5^3 = 125$	$(124, 21) = 1$	$(126, 21) = 21$
8	2	$8^1 = 8$	$(7, 21) = 7$	$(9, 21) = 3$
10	6	$10^3 = 1000$	$(999, 21) = 3$	$(1001, 21) = 7$
11	6	$11^3 = 1331$	$(1330, 21) = 7$	$(1332, 21) = 3$
13	2	$13^1 = 13$	$(12, 21) = 3$	$(14, 21) = 7$
16	3			
17	6	$17^3 = 4913$	$(4912, 21) = 1$	$(4914, 21) = 21$
19	6	$19^3 = 6859$	$(6858, 21) = 3$	$(6860, 21) = 7$
20	2	$20^1 = 20$	$(19, 21) = 1$	$(21, 21) = 21$

5 blanks, 6 hits, $w = \frac{6}{11}$

(b) $n = 437$ and $a = 94$
Solution: The period is $r = 22$ (calculation best by computer). It follows that

$$gcd\left(94^{11} - 1, 437\right) = 23;\ gcd\left(94^{11} + 1, 437\right) = 19;\ 23 \cdot 19 = 437. \tag{S.31}$$

2. We denote by w the probability that the conditions (S.3) are satisfied for the number $N = p_1^{\alpha_1} p_2^{\alpha 2} \dots p_m^{\alpha_m}$ (i.e. r is even and $a^{\frac{r}{2}} \mod N \neq 1$). Then the inequality $w \geq 1 - \frac{1}{2^{m-1}}$ applies. Particularly for $m = 2$ (the 'hardest' case), it follows that $w \geq 1 - \frac{1}{2^{m-1}} = \frac{1}{2}$. Check the findings for $N = 21$ and $N = 33$.

   The solution is found in Table S.2. As is seen, we have 5 blanks, 6 hits, $w = \frac{6}{11}$.
   The solution is found in Table S.3. As is seen, we have 9 blanks, 10 hits, $w = \frac{10}{19}$.
3. For which values of l is the sum

$$A = \left| \sum_{k=0}^{s-1} e^{2\pi i \frac{krl}{M}} \right|;\ l \in \mathbb{N}_0;\ r, M \in \mathbb{N} \tag{S.32}$$

   maximal?
   Solution: For $\frac{rl}{M} \in \mathbb{N}_0$, we have for each summand $e^{2\pi i \frac{krl}{M}} = 1$; it follows $A = s$. For $\frac{rl}{M} \notin \mathbb{N}_0$, each summand has magnitude 1, but there are different phases in dependence on k, such that the absolute value of the sum *must* be smaller than s.
4. Calculate

$$c\,(J, r, s; l) = \sum_{k=0}^{s-1} e^{2\pi i \frac{(J+kr)l}{M}}. \tag{S.33}$$

   Solution: We have

Table S.3 Solution of exercise 2 (b), $N = 33$

a	r	$a^{r/2}$	$gcd(a^{r/2} - 1, 15)$	$gcd(a^{r/2} + 1, 15)$
2	10	$2^5 = 32$	$(31, 33) = 1$	$(33, 33) = 33$
4	5			
5	10	$5^5 = 3125$	$(3124, 33) = 11$	$(3126, 33) = 3$
7	10	$7^5 = 16807$	$(16806, 33) = 3$	$(16808, 33) = 11$
8	10	$8^5 = 32768$	$(32767, 33) = 1$	$(32769, 33) = 33$
10	2	$10^1 = 10$	$(9, 33) = 3$	$(11, 33) = 11$
13	10	$13^5 = 371293$	$(371292, 33) = 3$	$(371294, 33) = 11$
14	10	$14^5 = 537824$	$(537823, 33) = 11$	$(537825, 33) = 3$
16	5			
17	10	$17^5 = 1419857$	$(1419856, 33) = 1$	$(1419858, 33) = 33$
19	10	$19^5 = 2476099$	$(2476098, 33) = 3$	$(2476100, 33) = 11$
20	10	$20^5 = 3200000$	$(3199999, 33) = 11$	$(3200001, 33) = 3$
23	2	$23^1 = 23$	$(22, 33) = 11$	$(24, 33) = 3$
25	5			
26	10	$26^5 = 11881376$	$(11881375, 33) = 11$	$(11881377, 33) = 3$
28	10	$28^5 = 17210368$	$(17210367, 33) = 3$	$(17210369, 33) = 11$
29	10	$29^5 = 20511149$	$(20511148, 33) = 1$	$(20511150, 33) = 33$
31	5			
32	2	$32^1 = 32$	$(31, 33) = 1$	$(33, 33) = 33$

9 blanks, 10 hits, $w = \frac{10}{19}$

$$c\left(J, r, s; l\right) = \sum_{k=0}^{s-1} e^{2\pi i \frac{(J+kr)l}{M}} = e^{2\pi i \frac{Jl}{M}} \sum_{k=0}^{s-1} \left(e^{2\pi i \frac{rl}{M}}\right)^k = e^{2\pi i \frac{Jl}{M}} \frac{1 - e^{2\pi i \frac{rl}{M}s}}{1 - e^{2\pi i \frac{rl}{M}}}. \tag{S.34}$$

With

$$e^{2\pi i \frac{Jl}{M}} \frac{1 - e^{2\pi i \frac{rl}{M}s}}{1 - e^{2\pi i \frac{rl}{M}}} = e^{2\pi i \frac{Jl}{M}} \frac{e^{\pi i \frac{rl}{M}s}}{e^{\pi i \frac{rl}{M}}} \frac{e^{-\pi i \frac{rl}{M}s} - e^{\pi i \frac{rl}{M}s}}{e^{-\pi i \frac{rl}{M}} - e^{\pi i \frac{rl}{M}}}$$
$$= e^{2\pi i \frac{Jl}{M}} e^{\pi i \frac{rl}{M}(s-1)} \frac{\sin \pi \frac{rl}{M}s}{\sin \pi \frac{rl}{M}} = e^{\pi i l \frac{2J + r(s-1)}{M}} \frac{\sin \pi \frac{rl}{M}s}{\sin \pi \frac{rl}{M}}, \tag{S.35}$$

it follows that

$$c\,(J, r, s; l) = \begin{cases} e^{\pi i l \frac{2J + r(s-1)}{M}} \frac{\sin \pi \frac{rl}{M}s}{\sin \pi \frac{rl}{M}} & ; \quad \frac{rl}{M} \notin \mathbb{N}_0 \\ e^{\pi i l \frac{2J}{M}} \cdot s & \quad \frac{rl}{M} \in \mathbb{N}_0. \end{cases} \tag{S.36}$$

The lower expression follows from the upper one by L'Hôpital's rule.
For the absolute value, we have the simple relationship

$$|c\,(J, r, s; l)| = \begin{cases} \left| \frac{\sin \pi \frac{rl}{M}s}{\sin \pi \frac{rl}{M}} \right| & ; \quad \frac{rl}{M} \notin \mathbb{N}_0 \\ s & \quad \frac{rl}{M} \in \mathbb{N}_0. \end{cases} \tag{S.37}$$

Note: From the previous problem, we know that $\left| \frac{\sin \pi \frac{rl}{M}s}{\sin \pi \frac{rl}{M}} \right| \leq s$. This will be shown in the next exercise by a different route.

5. Prove the inequality

$$|\sin (ny)| \leq n\,|\sin y|\,; \; n = 1, 2, 3... \tag{S.38}$$

Solution: We use induction. For $n = 1$, the inequality is evidently satisfied. We assume that it is also satisfied for n. Then we have for $n + 1$ according to the addition theorems for trigonometric functions:

$$|\sin ((n + 1)y)| = |\sin (ny) \cos y + \cos(ny) \sin y|\,. \tag{S.39}$$

We estimate the right side:

$$|\sin (ny) \cos y + \cos(ny) \sin y| \leq |\sin (ny) \cos y| + |\cos(ny) \sin y|\,. \tag{S.40}$$

For the first term on the right, we insert $|\sin (ny)| \leq n\,|\sin y|$; it follows that

$$|\sin ((n + 1)y)| \leq n\,|\sin y|\,|\cos y| + |\cos(ny)|\,|\sin y|\,. \tag{S.41}$$

On the right side, we factor out $|\sin y|$ and use $|\cos y| \leq 1$, $|\cos(ny)| \leq 1$. We then have

$$|\sin ((n + 1)y)| \leq (n + 1)\,|\sin y|\,, \tag{S.42}$$

and thus the statement is proven.

6. Formulate the single steps of the Shor algorithm explicitly for $N = 35$ and $a = 13$. Choose as in the text $M = 2^m = 128$.

Solution: From the first step, we obtain

$$|\varphi_1\rangle = \frac{1}{\sqrt{128}} \left(|0\rangle + |1\rangle + \cdots |127\rangle\right) |0\rangle . \tag{S.43}$$

Then we transform unitarily and find

$$|\varphi_2\rangle = \frac{1}{\sqrt{128}} \sum_{j=0}^{127} |j\rangle \left|13^j \bmod 35\right\rangle . \tag{S.44}$$

Because of

$$\begin{aligned} &13^0 \bmod 35 = 1;\ 13^1 \bmod 35 = 13; \\ &13^2 \bmod 35 = 29;\ 13^3 \bmod 35 = 27;\ 13^4 \bmod 35 = 1, \end{aligned} \tag{S.45}$$

it follows that $r = 4$. Thus we obtain in detail

$$|\varphi_2\rangle = \frac{|0\rangle |1\rangle + |1\rangle |13\rangle + |2\rangle |29\rangle + |3\rangle |27\rangle + |4\rangle |1\rangle + \cdots}{\sqrt{128}}. \tag{S.46}$$

Re-sorting gives

$$\begin{aligned} |\varphi_2\rangle = &\frac{|0\rangle + |4\rangle + |8\rangle + \cdots + |124\rangle}{\sqrt{128}} |1\rangle + \frac{|1\rangle + |5\rangle + |9\rangle + \cdots + |125\rangle}{\sqrt{128}} |13\rangle \\ &+ \frac{|2\rangle + |6\rangle + |10\rangle + \cdots + |126\rangle}{\sqrt{128}} |29\rangle + \frac{|3\rangle + |7\rangle + |11\rangle + \cdots + |127\rangle}{\sqrt{128}} |27\rangle . \end{aligned} \tag{S.47}$$

The QFT acting on the argument register

$$U_{QFT} = \frac{1}{\sqrt{128}} \sum_{l,k=0}^{127} e^{2\pi i \frac{lk}{128}} |k\rangle \langle l| \tag{S.48}$$

leads to the state $|\varphi_3\rangle$ (written in detail, i.e. without using $c\,(J, r, s; m)$):

$$\begin{aligned} |\varphi_3\rangle &= U_{QFT} |\varphi_2\rangle \\ &= \frac{1}{128}\left(\sum_{k=0}^{127} e^{2\pi i \frac{0k}{128}} |k\rangle + \sum_{k=0}^{127} e^{2\pi i \frac{4k}{128}} |k\rangle + \sum_{k=0}^{127} e^{2\pi i \frac{8k}{128}} |k\rangle + \cdots \right. \\ &\quad \left. + \sum_{k=0}^{127} e^{2\pi i \frac{124k}{128}} |k\rangle \right) |1\rangle \\ &\quad + \frac{1}{128}\left(\sum_{k=0}^{127} e^{2\pi i \frac{k}{128}} |k\rangle + \sum_{k=0}^{127} e^{2\pi i \frac{5k}{128}} |k\rangle + \sum_{k=0}^{127} e^{2\pi i \frac{9k}{128}} |k\rangle + \cdots \right. \end{aligned}$$

$$+ \sum_{k=0}^{127} e^{2\pi i \frac{125k}{128}} |k\rangle \Big) |13\rangle$$

$$+ \frac{1}{128}\Big(\sum_{k=0}^{127} e^{2\pi i \frac{2k}{128}} |k\rangle + \sum_{k=0}^{127} e^{2\pi i \frac{6k}{128}} |k\rangle + \sum_{k=0}^{127} e^{2\pi i \frac{10k}{128}} |k\rangle + \cdots$$

$$+ \sum_{k=0}^{127} e^{2\pi i \frac{126k}{128}} |k\rangle \Big) |29\rangle$$

$$+ \frac{1}{128}\Big(\sum_{k=0}^{127} e^{2\pi i \frac{3k}{128}} |k\rangle + \sum_{k=0}^{127} e^{2\pi i \frac{7k}{128}} |k\rangle + \sum_{k=0}^{127} e^{2\pi i \frac{11k}{128}} |k\rangle + \cdots$$

$$+ \sum_{k=0}^{127} e^{2\pi i \frac{127k}{128}} |k\rangle \Big) |27\rangle ; \tag{S.49}$$

and re-sorting gives

$$|\varphi_3\rangle = \frac{1}{128} \sum_{k=0}^{127} \left[e^{2\pi i \frac{0k}{128}} + e^{2\pi i \frac{4k}{128}} + e^{2\pi i \frac{8k}{128}} + \ldots + e^{2\pi i \frac{124k}{128}} \right] |k\rangle |1\rangle$$

$$+ \frac{1}{128} \sum_{k=0}^{127} e^{2\pi i \frac{k}{128}} \left[e^{2\pi i \frac{0k}{128}} + e^{2\pi i \frac{4k}{128}} + e^{2\pi i \frac{8k}{128}} + \ldots + e^{2\pi i \frac{124k}{128}} \right] |k\rangle |13\rangle$$

$$+ \frac{1}{128} \sum_{k=0}^{127} e^{2\pi i \frac{2k}{128}} \left[e^{2\pi i \frac{0k}{128}} + e^{2\pi i \frac{4k}{128}} + e^{2\pi i \frac{8k}{128}} + \ldots + e^{2\pi i \frac{124k}{128}} \right] |k\rangle |29\rangle$$

$$+ \frac{1}{128} \sum_{k=0}^{127} e^{2\pi i \frac{3k}{128}} \left[e^{2\pi i \frac{0k}{128}} + e^{2\pi i \frac{4k}{128}} + e^{2\pi i \frac{8k}{128}} + \ldots + e^{2\pi i \frac{124k}{128}} \right] |k\rangle |27\rangle . \tag{S.50}$$

The square brackets are nonzero only for $k = 0, 32, 64, 96$ (note that $k \le 127$), since (see also (S.35)) we have:

$$\sum_{l=0}^{31} e^{2\pi i \frac{4lk}{128}} = e^{\pi i k \frac{31}{32}} \cdot \frac{\sin \pi k}{\sin \frac{\pi k}{32}}$$

$$= 32 \cdot \left(\delta_{k,0} + \delta_{k,32} + \delta_{k,64} + \delta_{k,96} \right) . \tag{S.51}$$

Thus it follows that

$$|\varphi_3\rangle = \frac{1}{4} \sum_{k=0}^{127} e^{\pi i k \frac{31}{32}} \cdot \left(\delta_{k,0} + \delta_{k,32} + \delta_{k,64} + \delta_{k,96} \right) |k\rangle |1\rangle$$

$$+ \frac{1}{4} \sum_{k=0}^{127} e^{2\pi i \frac{k}{128}} \cdot \left(\delta_{k,0} + \delta_{k,32} + \delta_{k,64} + \delta_{k,96}\right) |k\rangle |13\rangle$$
$$+ \frac{1}{4} \sum_{k=0}^{127} e^{2\pi i \frac{2k}{128}} \cdot \left(\delta_{k,0} + \delta_{k,32} + \delta_{k,64} + \delta_{k,96}\right) |k\rangle |29\rangle$$
$$+ \frac{1}{4} \sum_{k=0}^{127} e^{2\pi i \frac{3k}{128}} \cdot \left(\delta_{k,0} + \delta_{k,32} + \delta_{k,64} + \delta_{k,96}\right) |k\rangle |27\rangle , \qquad \text{(S.52)}$$

or, explicitly,

$$|\varphi_3\rangle = \frac{|0\rangle + |32\rangle + |64\rangle + |96\rangle}{4} |1\rangle + \frac{|0\rangle + i\,|32\rangle - |64\rangle - i\,|96\rangle}{4} |13\rangle + \frac{|0\rangle - |32\rangle + |64\rangle - |96\rangle}{4} |29\rangle + \frac{|0\rangle - i\,|32\rangle - |64\rangle + i\,|96\rangle}{4} |27\rangle . \qquad \text{(S.53)}$$

Each measurement of the argument register yields one of the values 0, 32, 64, 96, and this independently of the value of the function register. The greatest common divisor of 32, 64, 96 is 32; hence the period is $r = \frac{128}{32} = 4$.

7. Work through the Shor algorithm for $N = 35$ and $a = 19$. Choose again $M = 128$.
Solution: In the first step, we have

$$|\varphi_1\rangle = \frac{1}{\sqrt{128}} \sum_{j=0}^{127} |j\rangle\, |0\rangle , \qquad \text{(S.54)}$$

and in the second:

$$|\varphi_2\rangle = \frac{1}{\sqrt{128}} \sum_{j=0}^{127} |j\rangle \left|19^j \bmod 35\right\rangle . \qquad \text{(S.55)}$$

The period is $r = 6$ due to

$19^0 \bmod 35 = 1$	$19^1 \bmod 35 = 19$	$19^2 \bmod 35 = 11$	$19^3 \bmod 35 = 34$
$19^4 \bmod 35 = 16$	$19^5 \bmod 35 = 24$	$19^6 \bmod 35 = 1$	$19^7 \bmod 35 = 19$

Re-sorting yields

$$|\varphi_2\rangle = \frac{1}{\sqrt{128}} \sum_J \sum_{k=0}^{s-1} |J + 6k\rangle \left|19^J \bmod 35\right\rangle \text{ with } s = 1 + \left[\frac{127 - J}{6}\right]. \tag{S.56}$$

The offset can take on the values $J = 0, 1, 2, 3, 4, 5$; correspondingly, we have $s = 22$ for $J = 0, 1$ and $s = 21$ for $J = 2, 3, 4, 5$.

In the third step, we transform the state $|\varphi_2\rangle$ via QFT to the state $|\varphi_3\rangle$:

$$|\varphi_3\rangle = \frac{1}{128} \sum_J \sum_{l=0}^{127} c(J, 6, s; l) |l\rangle \left|19^J \bmod 35\right\rangle \tag{S.57}$$

with

$$|c(J, 6, s; l)| = \begin{cases} \frac{\sin \pi \frac{3l}{64} s}{\sin \pi \frac{3l}{64}}; & \frac{3l}{64} \notin \mathbb{N}_0 \\ s & \frac{3l}{64} \in \mathbb{N}_0 \end{cases} ; \; s = \begin{cases} 22 \\ 21 \end{cases} \text{for } J = \begin{matrix} 0, 1 \\ 2, 3, 4, 5 \end{matrix} . \tag{S.58}$$

We obtain particularly large values if $\frac{3l}{64}$ is (at least approximately) an integer. We obtain an integer for $l = 0$ and $l = 64$, but $\frac{3l}{64}$ is also approximately an integer for $l = 21$ and 22, 42 and 43, 85 and 86, 106 and 107, and thus $|c(J, 6, s; l)|$ is also comparatively large. A numerical example for $s = 22$ illustrates this:

$l =$	40	41	42	43	44	45	$\cdots$	64
$\|c(J, 6, 22; l)\|^2 =$	3, 4	10, 1	72, 0	323, 0	22, 4	8, 1	$\cdots$	1024

If we calculate the period by $r \approx n \cdot \frac{M}{l}$, we find for the special l values;

$l =$	21	22	42	43	64	85	86	106	107
$\frac{r}{n} =$	6, 09	5, 82	3, 05	2, 98	2	1, 51	1, 49	1, 21	1, 20

From the last row, $r = 6$ results, so that we have determined the period by using the measurement data.

Appendix T
The Gleason Theorem

Gleason's theorem addresses the question of how we can define probabilities in quantum mechanics.

We assume that a system is in state $|\psi\rangle$; the density operator is given by $\rho = |\psi\rangle\langle\psi|$. Furthermore, we want to measure a property which we represent by the projection operator P. We denote by w_P the probability that the system in the state $|\psi\rangle$ yields the property associated with P in a measurement, or 'has' this property. Then, as we have derived in Chap. 22, w_P is given by

$$w_P = \langle P \rangle = tr\,(\rho P)\,. \tag{T.1}$$

The question is whether one can define the probabilities in a quite different way from (T.1). To state the question more precisely, we require as usual of the probabilities w_P the following properties:

$$\begin{gathered} 0 \le w_P \le 1 \text{ for all } P \text{ in the Hilbert space} \\ w_P\,(0) = 0;\;\; w_P\,(1) = 1;\;\; w_P\left(\sum_{i=1}^{\infty} P_i\right) = \sum_{i=1}^{\infty} w_P\,(P_i)\,. \end{gathered} \tag{T.2}$$

Under these conditions, the theorem of Gleason (1957) states that on a Hilbert space of dimension ≥ 3, the only possible probability measures are described by (T.1).[56] Since we want to use a formalism that applies for all dimensions, the restriction of the theorem to dim ≥ 3 is irrelevant, and we can say that all possible probability measures which can be defined in $\mathcal{H}$ are generated by the density operators of pure and mixed states.

It was soon recognized (by John Bell, among others) that Gleason's theorem is in conflict with our notions of realism. This is mainly because, according to Gleason's

[56] Perhaps this theorem thus provides a deeper reason for the special significance of density operators in quantum mechanics.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

theorem, the assignment of probabilities to all possible properties in a Hilbert space must be continuous, i.e. all vectors in the space must be mapped continuously into the interval [0, 1]. On the other hand, we understand the projectors P as a representation of yes-no observables, i.e. we can say of any property whether the system in fact possesses it or not. This results in a probability function that maps all P to 0 or 1, i.e. a discontinuous mapping. The Kochen–Specker theorem deals with this contradiction; see Chap. 27.

Appendix U
What is Real? Some Quotations

The questions 'What is real? What is reality? What is the nature of our knowledge of things?'[57] have preoccupied mankind from time immemorial; there are whole libraries on the subject. In the following, we have collected some quotes which are not meant as an attempt at a systematic exposition, but are simply incidental findings. The intention is rather to illustrate different positions—sometimes with wink of the eye.

The quotes are arranged by year of birth of the author (except the last one). Since the small collection has a rather casual character, and moreover a few quotes are probably incorrectly attributed to their 'authors', we dispense with detailed references.

1. 570 BC, Xenophanes, Greek philosopher,
   "And of course the clear and certain truth no man has seen nor will there be anyone who knows about the gods and what I say about all things. For even if, in the best case, one happened to speak just of what has been brought to pass, still he himself would not know. But opinion is allotted to all."
2. 460 BC, Democritus, Greek philosopher,
   "By convention sweet and by convention bitter, by convention hot, by convention cold, by convention color; but in reality atoms and void."
3. 428 BC, Plato, Greek philosopher,
   A precise quote is missing, but since Plato's allegory of the cave is so well known and fundamental, it should be briefly summarized. Some people are living in a cave. They can look only at a wall, and behind them a fire is burning. Items that

[57] Occasionally, the terms 'actuality' and 'reality' are used with different meanings. A distinction may boil down to the usage that *actuality* encompasses all objectively true statements, regardless of whether they are known or apparent to us at all, while *reality* includes any statements that we believe to be true (reality is what we perceive, actuality is what truly is). Another distinction is based on etymology: reality (Latin res = thing) refers to the materiality, actuality (to act) to the aspect of interaction or cause-effect. Since we want simply to illustrate here several different points of view and have no high philosophical aspirations, we accept the usage of common language which treats reality and actuality as synonyms.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

are carried past between the fire and the backs of the people cast their shadows on the wall. The people now know nothing about the items themselves, but only their shadows, so they take them for the ‘real’ world. Just like these people in the cave, we see only a semblance of true existence (i.e. the Platonic ideals); only philosophy can lead us to a perception of the ‘real’. (Perhaps Plato today would refer to the virtual world of television or computer screens instead of the shadows on a cave’s wall...). The influence of Plato on philosophy was immense for many centuries. In the words of Stephen Jay Gould: “The spirit of Plato dies hard. We have been unable to escape the philosophical tradition that what we can see and measure in the world is merely the superficial and imperfect representation of an underlying reality.”

4. 1303, Bridget of Sweden, Swedish mystic,
   “Although a blind man does not see it, the sun still shines clearly in splendor and beauty even while he is falling down the precipice.”
5. 1469, Niccolò Machiavelli, Italian politician and philosopher,
   “Men in general judge more from appearances than from reality. All men have eyes, but few have the gift of penetration.”
   “For the great majority of mankind are satisfied with appearances, as though they were realities, and are often even more influenced by the things that seem than by those that are.”
6. 1623, Blaise Pascal, French philosopher, physicist and mathematician,
   “Something incomprehensible is not for that reason less real.”
7. 1646, Gottfried Wilhelm Leibniz, German philosopher, scientist, diplomat, politician (‘the last polymath’),
   “Quite often a consideration of the nature of things is nothing but the knowledge of the nature of our minds and of these innate ideas, and there is no need to look for them outside oneself.”
8. 1685, George Berkeley, Irish theologian and philosopher,
   “*Esse est percipi*.” (“To be means to be perceived.”) Also quoted as “*Esse est percipi vel percipere*.” (“To be means to be perceived or to perceive.”). Only perceptions and perceivers really exist. The moon is not there when no one perceives it.
9. 1724, Immanuel Kant, German philosopher,
   “Hitherto it has been assumed that all our knowledge must conform to objects. But all attempts to extend our knowledge of objects by establishing something in regard to them a priori, by means of concepts, have, on this assumption, ended in failure. We must therefore test whether we might not have more success in the tasks of metaphysics, if we suppose that objects must conform to our knowledge. This would agree better with what is desired, namely, that it should be possible to have knowledge of objects a priori, determining something in regard to them prior to their being given. We should then proceed precisely on the lines of Copernicus’ primary hypothesis. Lacking satisfactory progress in explaining the movements of the heavenly bodies on the supposition that they all revolved around the spectator, he tested whether he might not have more success if he allowed the spectator to revolve and the stars to remain at rest.”

10. 1742, Georg Christoph Lichtenberg, German writer, aphorist, mathematician and the first German professor of experimental physics,
"Euler says in his letters upon various subjects in connection with natural science (Vol. II, p. 228), that there would be thunder and lightning just as well if there were no man present whom the lightning might strike. It is a very common expression, but I must confess that it has never been easy for me to comprehend it completely. It always seems to me as if the conception *being* were something derived from our thought, and thus, if there were no longer any sentient and thinking creatures, then there would be nothing more whatever. Although this sounds simpleminded, and although I would be laughed at if I said something like that publicly, I think yet that it is one of the greatest benefits, actually one of the strangest qualities of the human spirit, to be able to make such a conjecture."
11. 1770, Georg Wilhelm Friedrich Hegel, German philosopher,
"The spiritual alone is the real."
"Reason is the conscious certainty of being all reality."
"What is rational is real; and what is real is rational."
12. 1772, Novalis (Georg Philipp Friedrich Freiherr von Hardenberg), German poet and philosopher,
"Love is the highest reality—the deepest basis of everything."
13. 1821, Fyodor Mikhailovich Dostoyevsky, Russian writer,
"But does it matter whether it was a dream or reality, if the dream made known to me the truth?"
14. 1844, Friedrich Nietzsche, German philosopher,
"No, facts are precisely what there is not, only interpretations. We cannot establish any fact 'in itself'; perhaps it is folly to want to do such a thing. 'Everything is subjective', you say; but even this is interpretation."
15. 1871, Christian Morgenstern, German writer.
"For, he reasons pointedly, that which must not, can not be."
16. 1871, Marcel Proust, French writer,
"Reality is always the bait that lures us towards something unknown along a path that we can follow only a little way."
17. 1879, Albert Einstein, German-Swiss-American physicist,
"Reality is merely an illusion, albeit a very persistent one."
"As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality."
18. 1881, Pablo Picasso, Spanish painter,
"Everything you can imagine is real."
19. 1887, Marc Chagall, Russian-French painter,
"All our interior world is reality, and that, perhaps, more so than our apparent world."
20. 1887, Erwin Schrödinger, Austrian physicist,
"Reality is nothing more than a convenient fiction."
21. 1899, Alfred Hitchcock, British film director and producer,
"A glimpse into the world proves that horror is nothing other than reality."

22. 1889, Martin Heidegger, German philosopher,
“Higher than reality is potentiality.”
“Being essences (or happens, occurs) as presencing.”
23. 1901, Jacques Lacan, French psychoanalyst,
“If there is a notion of the real, it is extremely complex and, because of this, incomprehensible; it cannot be comprehended in a way that would make an All out of it.”
“The obviousness of reality, the obviousness of our sight which comprehends and takes up the world in terms of a logic of knowledge and of an understanding in pictures, is becoming questionable.”
24. 1903, Walker Evans, American photographer,
“Reality is not totally real.”
25. 1904, Salvador Dali, Spanish painter,
“One day it will have to be officially admitted that what we have christened reality is an even greater illusion than the world of dreams.”
26. 1906, Nelson Goodman, American philosopher,
“Truth, far from being a solemn and severe master, is a docile and obedient servant. The scientist who supposes that he is single-mindedly dedicated to the search for truth deceives himself.... He seeks system, simplicity, scope; and when satisfied on these scores he tailors truth to fit. He as much decrees as discovers the laws he sets forth, as much designs as discerns the patterns he delineates.”
27. 1914, Arno Schmidt, German writer,
“The > Real World <? is, in truth, only the caricature of our great novels!” (Schmidt used a very personal orthography; in German, the quote reads: “Die > Wirkliche Welt <? : ist, in Wahrheit, nur die Karikatur unsrer Großn Romane!”).
28. 1921, Paul Watzlawick, Austrian-American psychiatrist,
“The belief that one’s own view of reality is the only reality is the most dangerous of all delusions.”
29. 1928, Philip K. Dick, American SF-writer,
“Reality is that which, when you stop believing in it, doesn’t go away.”
30. 1928, Robert M. Pirsig, American writer and philosopher,
“Laws of nature are human inventions…the world has no existence whatsoever outside the human imagination.”
31. 1929, Audrey Hepburn, British actress and humanitarian,
“Anyone who does not believe in miracles is not a realist.”
32. 1929, Jean Baudrillard, French philosopher and sociologist,
“For the world was not created in order to understand it. It does not care about knowledge. Maybe it was even created to be not understood. Knowledge is indeed part of the world, but only as a total illusion. That’s what I find interesting, because it means that the mind is only part of a whole, and that there is no interpretation for this whole thing. ... Inside this world there is definitely a knowledge- and thought system that produces something like truth- and reality effects. But I think it’s important that philosophy has always in mind this radical uncertainty and illusion. One must beware of the truth.”

33. 1931, Roger Penrose, British mathematician and physicist,
"It is my opinion that our present picture of physical reality, particularly in relation to the nature of time, is due for a grand shake up—even greater, perhaps, than that which has already been provided by present-day relativity and quantum mechanics."
34. 1935, David Mermin, American physicist,
"We know that the moon is demonstrably not there when nobody looks."
35. 1935, Woody Allen, American screenwriter, director, actor, comedian, author,
"I hate reality, but it's still the best place to get a good steak."
36. 1940, John Lennon, British musician,
"Nothing is real and nothing to get hung about. Strawberry Fields forever."
37. 1944, Yves Michaud, French philosopher,
"What we call reality is an unsatisfactory system of a small number of sensory experiences, of ill-founded beliefs and superficially perceived images."
38. 1945, Richard Tarnas, American philosopher,
"The world is in some essential sense a construct. Human knowledge is radically interpretive. There are no perspective-independent facts. Every act of perception and cognition is contingent, mediated, situated, contextual, theory-soaked. Human language cannot establish its ground in an independent reality."
39. 2010, Christian Lange, Nils Ohlsen (eds.): *Realism—The Adventure of Reality*, exhibition catalog for the exhibition at the Kunsthalle in Emden, Germany; January through June, 2010,
"Based on the paradigm of cultural studies of a 'radical constructivism' and on the theory of the collective reality production holds today, that reality is not objectively evident. ... One knows that the perception of people can hardly distinguish between the subjectivity in the phenomenal and objective reality. ... Obviously, the knowledge that reality can and must be continuously challenged contains just the prerequisite for the efforts to get, again and again, a new and most accurate idea of it."

Appendix V
Remarks on Some Interpretations of Quantum Mechanics

We consider in this appendix four interpretations in more detail than in Chap. 28, namely the Bohmian interpretation, the many-worlds interpretation, consistent histories and, finally, the Ghirardi-Rimini-Weber theory.

V.1 Bohmian Interpretation

This is not only the currently effectively unique interpretation of quantum mechanics based on hidden variables, but surely also the most thoroughly elaborated one. The basic idea is to consider the wavefunction and the associated particle as two different, separately existing and real objects. It is also given other names such as de Broglie-Bohm theory, pilot-wave theory, Bohmian mechanics, causal interpretation, etc.

V.1.1 Sketch of the Formalism

We begin with the SEq:

$$i\hbar\dot{\Psi} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi. \tag{V.1}$$

Since the wavefunction and probability density are related by $\rho = |\Psi|^2$, we apply the *ansatz*

$$\Psi = \rho^{\frac{1}{2}} e^{i\frac{S}{\hbar}};\ \rho, S \in \mathbb{R}, \tag{V.2}$$

so that S has the dimensions of an action. We insert (V.2) into the SEq and separate with respect to real part and imaginary part. This yields:

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

$$i\hbar\frac{1}{2}\rho^{-\frac{1}{2}}\dot{\rho} = -\frac{\hbar}{2m}\left[\rho^{-\frac{1}{2}}\nabla\rho\cdot i\nabla S + \rho^{\frac{1}{2}}i\nabla^2 S\right]$$
$$-\rho^{\frac{1}{2}}\dot{S} = -\frac{\hbar^2}{2m}\left[-\frac{1}{4}\rho^{-\frac{3}{2}}(\nabla\rho)^2 + \frac{1}{2}\rho^{-\frac{1}{2}}\nabla^2\rho - \rho^{\frac{1}{2}}\frac{1}{\hbar^2}(\nabla S)^2\right] + V\rho^{\frac{1}{2}} \quad \text{(V.3)}$$

or

$$\begin{aligned}\dot{\rho} + \frac{1}{m}\left[\nabla\rho\cdot\nabla S + \rho\nabla^2 S\right] &= \dot{\rho} + \nabla\left(\rho\cdot\frac{\nabla S}{m}\right) = 0\\ \dot{S} + \frac{1}{2m}(\nabla S)^2 + V - \frac{\hbar^2}{4m}\left[\rho^{-1}\nabla^2\rho - \frac{1}{2}\rho^{-2}(\nabla\rho)^2\right] &= 0.\end{aligned} \quad \text{(V.4)}$$

For $\hbar \to 0$ (i.e. in the classical limit), S is a solution of the Hamilton-Jacobi equation[58]; hence, the term ∇S can be interpreted as the momentum $\mathbf{p}$, or $\frac{\nabla S}{m}$ as $\mathbf{v}$, i.e. as the velocity of a point particle.[59]

This interpretation can be maintained for $\hbar \neq 0$. The first equation in (V.4) then reads $\dot{\rho} + \nabla(\rho\cdot\mathbf{v}) = 0$, which is none other than the conservation of probability. The second equation can be further understood as the Hamilton-Jacobi equation, but, in addition to the 'standard' potential V, it contains a *quantum potential* W, with

$$W = -\frac{\hbar^2}{4m}\left[\rho^{-1}\nabla^2\rho - \frac{1}{2}\rho^{-2}(\nabla\rho)^2\right] = -\frac{\hbar^2}{2m}\frac{\nabla^2\rho^{\frac{1}{2}}}{\rho^{\frac{1}{2}}} = -\frac{\hbar^2}{2m}\frac{\nabla^2|\psi|}{|\psi|}. \quad \text{(V.5)}$$

With this, (V.4) can be rewritten as

$$\begin{aligned}\dot{\rho} + \nabla\left(\rho\cdot\tfrac{\nabla S}{m}\right) &= \dot{\rho} + \nabla\left(\rho\cdot\tfrac{\mathbf{p}}{m}\right) = 0\\ \dot{S} + \tfrac{(\nabla S)^2}{2m} + V + W &= \dot{S} + \tfrac{\mathbf{p}^2}{2m} + V + W = 0.\end{aligned} \quad \text{(V.6)}$$

Here, it is important to note that the point particle experiences not only the classical interaction V, but also the quantum potential W.

In principle, one can now obtain the Hamilton-Jacobi function S by integrating equation (V.6). The speed of the point particle follows from this in terms of $\mathbf{v} = \frac{\nabla S}{m}$, and its trajectory from $\mathbf{x} = \int \mathbf{v}\, dt$. However, we do not know the initial value; the integral over the velocity gives a family (ensemble) of possible trajectories. Below, we will go through the procedure for the motion of a free particle as an example.

But first, some general remarks about the Bohmian interpretation: We see that the wavefunction ψ plays a double role. It supplies the information about the most probable position of the particle in terms of $\rho = |\Psi|^2$, and on the other hand, it affects the particle through the quantum potential W. In this interpretation, the physical state of a particle is completely defined not by the wavefunction alone, but by the combination of the wavefunction and the particle's position.

[58] This of course holds true only if the terms $\rho^{-1}\nabla^2\rho - \frac{1}{2}\rho^{-2}(\nabla\rho)^2$ do not vanish as $1/\hbar^2$, or vanish more slowly.

[59] Here, one in fact considers a point particle with well-defined position and velocity, and not a quantum object.

Since the wavefunction (or at least its absolute square) exerts a force on the particles by means of W, we have to consider ψ (or at least $|\psi|$) as the mathematical representation of a real field.[60] Of course, the coordinates of the particles are considered a priori as real; but they are not observable and represent the hidden variables in this interpretation.

Thus, the particles move along well-defined trajectories, which we do not know, but about which we can make probability statements with the help of ρ. Only the particles are relevant to the measurement; the wavefunction acts as a (somewhat nebulous) 'guiding field'.

In this interpretation, there is no collapse due to the measurement; instead, measurement simply means reducing our ignorance—after the measurement, we know on which trajectory the particle is moving (on which it was moving even before the measurement, but then we did not know this).

We note that the interpretation is nonlocal due to the occurrence of the quantum potential. We can see this easily if we imagine making changes on just one particle in a multi-particle system. Then its wavefunction ψ and quantum potential W change instantaneously, and hence the trajectories of all the other particles also must change.[61]

V.1.2 Example: Free Motion

For the sake of illustration, we consider the free motion of a particle. With the wavefunction

$$\psi = Ne^{i(kx-\omega t)} \tag{V.7}$$

(N is a suitable normalization factor), we obtain

$$S = \hbar\,(kx - \omega t)\,. \tag{V.8}$$

Evidently, the quantum potential W vanishes, and the Bohmian momentum p is given by

$$p = \frac{\mathrm{d}}{\mathrm{d}x}S = \hbar k. \tag{V.9}$$

With this and $p = m\dot{x}$, the family of trajectories follows:

$$x = \frac{\hbar k}{m}t + x_0, \tag{V.10}$$

[60]Where this field comes from and what its physical cause is are left unspecified.

[61]According to Bell, the term 'hidden variables', which is commonly used for variables that are intended to supplement the quantum-mechanical description (using wavefunctions), is misleading here, since on the contrary, the wavefunction is 'hidden'.

whereby we do not know the specific initial value x_0 which applies to the system under consideration. Since the probability density (in the distribution sense) is given by $|N|^2$, all initial positions are equally probable.

V.1.3 Conclusions

The Bohmian interpretation yields the same results as the usual quantum mechanics. Moreover, it offers no advantages, either formally or computationally: although it has recourse to classical ideas, it cannot avoid phenomena such as non-locality, in contradiction to common sense. Hence a striking advantage of this approach is not apparent. In addition, there are other difficulties not mentioned here, including among others the extension to multi-particle systems or to relativistic velocities. A further main point of criticism is the asymmetry between position and momentum. Quantum mechanics can be formulated, for example, either in position or in momentum space. This is not true for the Bohmian interpretation; it depends largely on the position representation.

The attitude of the scientific community towards the Bohmian interpretation differs greatly (as with all interpretations). While many do not see it any longer as a serious explanation of quantum mechanics, others do not share this view; for example, the group 'Bohmian Mechanics' at the University of Munich.[62]

V.2 The Many-Worlds Interpretation

The many-worlds interpretation (MWI) is based on the unmodified SEq. In contrast to the standard interpretation, it assumes that all of the changes that the SEq describes over time will be realized; it is therefore a strictly deterministic theory.

We illustrate this fact with the help of a photon of unknown polarization which is measured with respect to its horizontal/vertical polarization. The basis states of the photon are $|h\rangle$ and $|v\rangle$; the measuring apparatus has the states $|M_h\rangle$ and $|M_v\rangle$. Then the total state is:

$$|\psi\rangle = c_h\,|h\rangle\,|M_h\rangle + c_v\,|v\rangle\,|M_v\rangle\,;\;\; |c_h|^2 + |c_v|^2 = 1. \tag{V.11}$$

In the standard interpretation, this means that we measure the state $|i\rangle$ with a probability of $|c_i|^2$, where the measuring apparatus is in the state $|M_i\rangle$ $(i = h, v)$. The MWI, however, assumes that both terms on the right side of (V.11) describe something that really exists. So there is no state reduction and no reference to probabilities occurring in measurements. A common view is that the universe branches out into a number of different parallel worlds; in our example into two universes, one with

---

[62] http://www.mathematik.uni-muenchen.de/~bohmmech/.

$|h\rangle\,|M_h\rangle$ and the second with $|v\rangle\,|M_v\rangle$. Thus, the concept of 'measurement' plays no fundamental role, which is the reason why the MWI is currently enjoying some popularity, especially with quantum cosmologists.

However, there are some problems with the MWI connected with the concept of probability. Let us consider again the state (V.11). Both 'worlds', i.e. $|h\rangle\,|M_h\rangle$ and $|v\rangle\,|M_v\rangle$, exist after the split. What then is the significance of the coefficients c_h and c_v? It is clear is that $|c_h|^2$ and $|c_v|^2$ cannot be interpreted as probabilities for the occurrence of one or another world, since *both* are realized. There are different explanations,[63] for example the many-minds interpretation, according to which each conscious being has available a continuum of states of consciousness, representing the branched worlds.

In science fiction, parallel universes allow for all sorts of spectacular activities. But in fact, and here the various offshoots of the MWI agree, we experience nothing of these branches (and thus cannot visit or communicate with parallel universes), because in each branch of the state vector there is a perfect correlation between our memory states and the other events which have occurred.

Another problem of the MWI is that the decomposition or branching is not always unique. We consider a system of two electrons with total spin zero. The antisymmetric (and entangled) total state is

$$|\psi\rangle = \frac{|\uparrow\rangle\,|\downarrow\rangle - |\downarrow\rangle\,|\uparrow\rangle}{\sqrt{2}}, \tag{V.12}$$

where $|\uparrow\rangle$ and $|\downarrow\rangle$ are the eigenstates (referred to the z axis) for the eigenvalues $+\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$. In the MWI, this state describes a split into two branches, one with $|\uparrow\rangle\,|\downarrow\rangle$ and one with $|\downarrow\rangle\,|\uparrow\rangle$.

Instead of the z axis, we could just as well refer to the x axis. The spin matrices are

$$S_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix};\; S_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{V.13}$$

The eigenvectors for the eigenvalues $+\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$ are given for S_x by

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix};\; |\leftarrow\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \tag{V.14}$$

and for S_z by

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix};\; |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{V.15}$$

[63] In any case, the question cannot be decided by an outside observer who can determine, so to speak, the 'weight' of the individual worlds— simply because there is not such an 'outsider's perspective' in the MWI (in other words, because for our universe, according to the common belief, there is neither an 'outside' nor a 'before', i.e. neither anteriority nor exteriority).

This us allows to convert the state (V.12), formulated in the z basis, into the x basis. It follows that

$$|\psi\rangle = \frac{|\leftarrow\rangle\,|\rightarrow\rangle - |\rightarrow\rangle\,|\leftarrow\rangle}{\sqrt{2}}. \tag{V.16}$$

We see that the total states (V.12) and (V.16) are the same, but not the individual states on the right-hand sides. This means that also the possible ramifications are different. The question arises: Which is the correct splitting, or what is the special basis in the Hilbert space which is related to the branching? We discussed this issue in Chap. 24 in the context of decoherence (non-uniqueness of a biorthonormal decomposition for the same coefficients). Similarly, we would have to assume, if we demand uniqueness, that by the interaction with the environment, a pointer variable is selected that determines which one of the two options (V.12) and (V.16) is realized. Apparently, however, this question can not be answered satisfactorily at present.

All in all, there are several points in the MWI which seem still not completely resolved on a quite fundamental level; perhaps the naive notion of a constantly-branching universe can be improved upon.

V.3 Consistent Histories

V.3.1 Definitions

The term *quantum history* essentially means 'time-ordered sequence of quantum events'. The term 'event' is quite flexible; events can be e.g. wavefunctions or properties such as position or momentum or others. We will represent each event (seen as a property of the system) by a projection operator F. In this section, we use subscript indices exclusively to indicate different times; for other distinctions, we choose accordingly superscript indices. For a given time sequence $t_1 < t_2 < t_f$, a history is characterized by a sequence of projectors $(F_1, F_2, \ldots, F_f)$, one projector for each time.

One can imagine a history as a stroboscopic recording of a (possibly continuous) process, but the intervals need not be constant and the events considered must not be the same at each time step. As an example, we consider a harmonic oscillator with eigenstates $\left\{\left|\varphi^i\right\rangle, i = 1, 2, \ldots\right\}$. At the first time step, we project onto the subspaces 1 and 2, at the second onto the subspace 3 and at the third, we measure whether the system is in the region $a \leq x \leq b$:

$$F_1 = \left|\varphi^1\right\rangle\left\langle\varphi^1\right| + \left|\varphi^2\right\rangle\left\langle\varphi^2\right|;\ F_2 = \left|\varphi^3\right\rangle\left\langle\varphi^3\right|;\ F_3 = X^{a,b}. \tag{V.17}$$

The crucial point is that the successive events need not be linked via the SEq, i.e. one can also incorporate e.g. stochastic evolutions. However, the SEq (or equivalently the time-evolution operator $T(t', t)$) may be used for the calculation of probabilities.

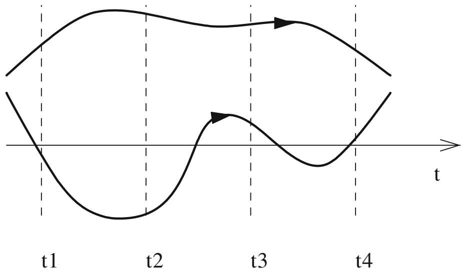


Fig. V.1 Schematic representation of different histories

We can now define a *history Hilbert space* $\breve{\mathcal{H}}$ by

$$\breve{\mathcal{H}} = \mathcal{H}_1 \odot \mathcal{H}_2 \odot \cdots \mathcal{H}_f, \tag{V.18}$$

where $\mathcal{H}_j$ is a copy of the Hilbert space $\mathcal{H}$ for each time point t_j, which describes the system at a fixed time. Like $\otimes$, the symbol $\odot$ denotes a tensor product; the difference is that $\otimes$ couples Hilbert spaces at the same times, and $\odot$ at different times. In the space $\breve{\mathcal{H}}$, the history $(F_1, F_2, \ldots, F_f)$ can be represented by the tensor product

$$Y = F_1 \odot F_2 \odot \cdots F_f. \tag{V.19}$$

This formulation means 'F_1 at t_1, F_2 at t_2, …, F_f at t_f' ; at other times, it provides no information. Since each F_i is a projector, this also applies to Y.

Between two points in time, there may be different histories, as indicated in Fig. V.1. One can now assign a probability to each history, so that one can distinguish between more or less probable histories. To this end, one defines an operator $K(Y)$ (chain operator):

$$K(Y) = F_f T(t_f, t_{f-1}) F_{f-1} T(t_{f-1}, t_{f-2}) \cdots T(t_1, t_0) F_0 \tag{V.20}$$

where the operators $T(t_j, t_{j'})$ are the corresponding time evolution operators which convey the system between the points $t_{j'}$ and t_j. The probability for the history (V.19) is determined by

$$W(Y) = tr\left(K^{\dagger}(Y) K(Y)\right). \tag{V.21}$$

Some properties of a history can be read directly off these formulations. For instance, W vanishes for $F_f T(t_f, t_{f-1}) F_{f-1} = 0$, i.e. for a contradictory history.

As an example, we consider projections onto pure states, e.g. $|a\rangle$, $|b\rangle$ and $|c\rangle$. It follows that:

$$Y = |a_0\rangle \langle a_0| \odot |b_1\rangle \langle b_1| \odot |c_2\rangle \langle c_2|, \tag{V.22}$$

and the chain operator

$$\begin{aligned} K(Y) &= |c_2\rangle \langle c_2| T(t_2, t_1) |b_1\rangle \langle b_1| T(t_1, t_0) |a_0\rangle \langle a_0| \\ &= \langle c_2| T(t_2, t_1) |b_1\rangle \langle b_1| T(t_1, t_0) |a_0\rangle \cdot |c_2\rangle \langle a_0| \end{aligned} \tag{V.23}$$

is a product of complex numbers (transition amplitudes) with the operator $|c_2\rangle \langle a_0|$.

One can join different histories into a family, if they meet certain *consistency conditions*. Intuitively, the conditions mean that the probabilities of various histories are additive. For two histories Y^a and Y^b, this is the case for

$$tr\left(K^\dagger(Y^a)K(Y^b)\right) = 0 \text{ for } a \neq b. \tag{V.24}$$

Of course, the probabilities of all observed histories have to add up to give 1. Such a set of histories is called a *consistent family of histories* or framework. In essence, this means that a consistent family of histories consists of different, mutually exclusive histories.

The consistent histories approach now assumes that for the description of a measurement with certain results, a framework must be used in which these results are included, and that the framework must also include the measurement properties at a time before the measurement takes place. Within this consistent family, there are no contradictions, no paradoxes. These occur only when histories from different (i.e. incompatible) families are compared.

A final note: The consistency conditions (V.24) are also called *decoherence conditions*, and accordingly, the expressions 'decoherent family' or 'decoherent set' are occasionally used to denote a consistent family. The use of 'decoherence' in this context is perhaps a somewhat confusing example of language, since it is *not* directly related to the phenomenon of decoherence as a flow of information into the environment (which we treated in Chap. 24; although they may have something to do with each other). Thus, a consistent family of histories can become inconsistent for an isolated system if the environment is taken into account, and *vice versa*, an inconsistent family can become a consistent one.

V.3.2 A Simple Example

For a very simple illustration of the basic ideas, we consider a Mach–Zehnder interferometer, where the beam splitter BS2 is mobile and can be moved out of/into the beam path; see Fig. V.2.

We denote the initial state of the photon by $|a\rangle$ (horizontal incidence), the upper and lower beam paths by $|o\rangle$ and $|u\rangle$, and the superposition state by $|s\rangle$; apart from that, we use the results of Chap. 6, Vol. 1. We write the total state as $|\varphi\rangle\, |AB\rangle$, i.e. as the tensor product of the Hilbert spaces of the photon $|\varphi\rangle$ and the two detectors $|AB\rangle$. If the photon triggers a detector, for example A, it is absorbed and vanishes; we write this as $|\varphi\rangle\, |AB\rangle \rightarrow |A^*B\rangle$.

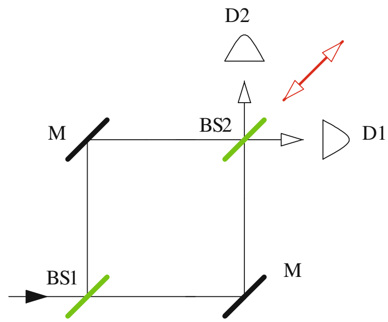


Fig. V.2 Mach–Zehnder interferometer used for the simple example illustrating consistent histories

(1) We first consider the situation where the beam splitter BS2 is not in the path. The incoming photon is incident on the first beam splitter, after which it triggers either detector D1 or D2, and thus has to go through the lower or the upper arm. This we can write as

$$|a\rangle\,|D1D2\rangle \to \begin{pmatrix} |o\rangle\,|D1D2\rangle \to |D1^{*}D2\rangle\,;\;\; W=\frac{1}{2} \\ |u\rangle\,|D_1D_2\rangle \to |D1D2^{*}\rangle\,;\;\; W=\frac{1}{2} \end{pmatrix}. \tag{V.25}$$

Thus, we have formulated a consistent family of histories; the probabilities of the two family members add up to 1. The two paths (and thus the displays of the detectors) are mutually exclusive.

(2) We now bring the beam splitter BS2 into the beam path. Then we have the following family of consistent histories.

$$|a\rangle\,|D1D2\rangle \to \begin{pmatrix} |s\rangle\,|D1D2\rangle \to |D1^{*}D2\rangle\,;\;\; W=1 \\ |s\rangle\,|D1D2\rangle \to |D1D2^{*}\rangle\,;\;\; W=0 \end{pmatrix}. \tag{V.26}$$

The photon initially strikes the first beam splitter, then passes into the superposition state, meets the second beam splitter and triggers (at least in principle) D1/D2. As we have shown in Chap. 6, Vol. 1, for a horizontally incident photon, the upper history has the probability 1, the lower one 0; only D1 responds.

Thus we have two different and mutually-exclusive families of histories—no member of one family can belong to the other family. Accordingly, questions that are useful for one family can be meaningless for the other, for example the question as to which path the photon has followed.

V.3.3 Conclusions

Consistent histories is an approach that characterizes the physics of a closed system by a discrete time series of properties that are represented by projectors in Hilbert space. The successive events need not be connected by the SEq; the evolution can also be stochastic. Consistency conditions ensure a consistent physical interpretation of the events.

Since here the SEq is not *the* general governing principle of the time evolution, but one of several possible ones, this interpretation avoids many problems that are linked to the concept of measurement—questions about the classical nature of the measuring apparatus, the mechanism of the collapse of the wavefunction, etc., do not even arise. On the other hand, this approach also cannot do without the SEq if probabilities for different histories are to be calculated. And fundamental questions remain unanswered here, also—for example whether objective chance exists or how nature selects out of the family of histories that one history which we actually measure.

There is a comprehensive primary and secondary literature on consistent histories.[64] An introduction is given by e.g. Robert B. Griffiths, *Consistent Quantum Theory* (2003). The online version of the book can be found at http://quantum.phys.cmu.edu/CQT/.

V.4 Ghirardi-Rimini-Weber

This approach goes beyond a simple interpretation, since the dynamics is modified. The aim is to suppress superpositions of states which are located in widely-separated regions (i.e. macroscopically separated) through an additional term in the Hamiltonian. In other words, the modified dynamics conditions are constructed in such a way that they cause the transition from a pure state into a statistical mixture under macroscopic conditions. Thus it is clear that the starting point is not the SEq for pure states, but the von Neumann equation for the density operator:

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) = -\frac{i}{\hbar}\left[H, \rho(t)\right]. \tag{V.27}$$

In this equation, a term is inserted which corresponds to a localization process:

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho = -\frac{i}{\hbar}\left[H, \rho\right] - \lambda\left(\rho - T\left[\rho\right]\right), \tag{V.28}$$

with

$$T\left[\rho\right] = \sqrt{\frac{\alpha}{\pi}} \int\limits_{-\infty}^{\infty} \left(e^{-\frac{\alpha}{2}(q-x)^2} \rho e^{-\frac{\alpha}{2}(q-x)^2}\right) \mathrm{d}x, \tag{V.29}$$

where q is the position operator. The quantity λ describes the frequency of the localization process and $\frac{1}{\sqrt{\alpha}}$ corresponds to the distance after which the linear superposition changes into a statistical mixture (localization distance).

[64]The approach is applied e.g. also in quantum cosmology.

Intuitively, this term means that the quantum object is subject to random localization processes around its approximate position, at random times (Poisson distribution with main frequency λ). Formally, this localization can be summarized as[65]:

$$\psi(q) \to \psi_x(q) = \frac{\varphi_x(q)}{\|\varphi_x(q)\|};\ \varphi_x(q) = \left(\frac{\alpha}{\pi}\right)^{\frac{3}{4}} e^{-\frac{\alpha}{2}(q-x)^2}\psi(q). \tag{V.30}$$

If $\psi(q)$ is a wave packet centered around q_0, then $\varphi_x(q)$ in general does not yield noteworthy contributions for $x \neq q_0$.

We mention here only that the localization process is proportional to the number of quantum objects involved—the larger the number, the faster the total system becomes localized.

The quantities λ and α are not physically defined, but are freely adjustable parameters; they are chosen in such a way that they meet the predetermined requirements as well as possible. Typical values are $\lambda \approx 10^{-16}\ \text{s}^{-1}$ and $\frac{1}{\sqrt{\alpha}} \approx 10^{-7}$ m. With these values, one obtains localization times of e.g. 10^{16} s ($\approx 3 \cdot 10^7$ a) for microscopic and 10^{-7} s for macroscopic systems.

The GRW approach thus alters standard quantum mechanics so that it yields almost unchanged results for microscopic systems, but ensures the decay of superpositions for macroscopic systems.

Among other things, a weak point of this approach can be seen in the fact that the localization process is postulated ad hoc; its physical origin (if it even exists) is not considered. So it is rather an empirically- and phenomenologically-oriented model with free model parameters which are adapted to fit measurements.

[65]The generalization to N quantum objects in three-dimensional space, if the ith quantum object is located at $\mathbf{q}_i$:

$$\psi(\mathbf{q}_1, \ldots, \mathbf{q}_1) \to \psi_x(\mathbf{q}_1, \ldots, \mathbf{q}_N) = \frac{\varphi_x(\mathbf{q}_1,\ldots,\mathbf{q}_N)}{\|\varphi_x(\mathbf{q}_1,\ldots,\mathbf{q}_N)\|}$$
$$\varphi_x(\mathbf{q}_1, \ldots, \mathbf{q}_N) = \left(\frac{\alpha}{\pi}\right)^{\frac{3}{4}} e^{-\frac{\alpha}{2}(\mathbf{q}_i-\mathbf{x})^2}\psi(\mathbf{q}_1, \ldots, \mathbf{q}_N).$$

Appendix W
Elements of Quantum Field Theory

W.1 Foreword

In Volume 1 we discussed some topics of relativistic quantum mechanics (RQM). In this volume we want to provide a glimpse into relativistic *quantum field theory* (QFT).

Why QFT? The subjects of quantum mechanics (QM) are characterized by three properties: (1) They are *single* particles, e.g. electrons. (2) They retain their *identity* (an electron remains an electron remains an electron) (3) They behave *non-relativistically*. RQM, which we have dealt with in Volume 1, extends the theory into the relativistic domain, thereby attempting to retain properties (1) and (2). But as we have seen with respect to property (1), there are no relativistic single-particle theories. In addition, properties (1) and (2) mean that there is no way to treat processes like the decay of elementary particles or reactions like $A + B \to C + D + E$. And finally, we have found quite problematic negative energies which can not be interpreted meaningfully in the framework of these equations (Klein–Gordon and Dirac equation). To attack all these questions, an advanced theory is required, and as we will see, QFT will resolve the issue.

Also in another regard, property (1) is a limitation, since it refers to individual particles, but not to extended objects such as, e.g. a string. How could one quantize such objects? A mass point has the parameters $x(t)$ and $\dot{x}(t)$ or $p(t)$ (here x is a Lagrange coordinate, 'moving in the river with the mass point'). If we combine many mass points $x_i(t)$ to form e.g. a string, we speak no longer of individual coordinates. Instead, we have a field variable $\varphi(x, t)$, e.g. the amplitude of the string (here x is an Euler coordinate, 'sitting on the bank of the river'). Thus, instead of the label i we have now the label x, and instead of quantizing x of the mass point (called 1. quantization), we now have to quantize the field φ (called 2. quantization).[66] As the name quantum *field* theory implies, QFT is the right candidate for this task.

[66]The naming 1. and 2. quantization is a little bit unfortunate but established.

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

The formalism relies very much on classical field theory (see outline in Appendix T, Vol. 1). A central term is the Lagrange density which allows for the determination of the conjugated momentum density and the Hamiltonian density. QFT starts from the equations of RQM, e.g. the Klein–Gordon and the Dirac equation. In principle one can also formulate a nonrelativistic quantum field theory, starting from the Schrödinger equation. Indeed, there are applications for this nonrelativistic field theory. Nevertheless, the name QFT means almost exclusively *relativistic* quantum field theory.

Another essential component of QFT is the formalism of the harmonic oscillator as developed in Vol. 2 Chap. 18. Actually, the quantization procedure for fields is formulated by means of creation and annihilation operators and uses terms like number operator and so on.

Among the advantages of QFT are the following: (1) Electrons and positrons have equal status; we do not have an electron sitting on an infinite sea of positrons anymore. (2) We get rid off negative energies; QFT knows positive energies only. (3) We can consider processes with say n particles in the incoming channel and m particles in the outgoing channel. (4) With quantum electrodynamics, we have the most stringently proven theory in physics. However, we should also mention that there exist structural problems with infinities in QFT. And on the technical level, QFT is sometimes more demanding than QM or RQM.

The content plan is as follows: First we show by means of a toy example the main ideas which are underlying the quantization of fields. This is followed by three sections in which we quantize the Klein–Gordon, Dirac and radiation field. As we will see, there occur infinities; the way to handle this problem is discussed in the section 'Operator ordering'. In the last sections, we formulate the theory for interacting fields using the example of quantum electrodynamics (QED). We define the S-matrix and its approximation by first order and second order terms. On the basis of first order terms, we address among others Feynman diagrams. In order to calculate the second order terms, we define Feynman propagators and present the Wick theorem. Finally, we consider the second order term of the S-matrix, exemplarily treating in more detail Bhabha and Møller scattering.

QFT is a very extensive topic. There are certainly quite different opinions about how to present it on a few pages as we try to do here. Be that as it may, in view of the space limitation we can not build here a royal road into the realm of quantum field theory, but can only pave the way with a few stepping stones. A lot of questions and issues cannot even be mentioned, and in the discussed topics there will be inevitably some gaps. The reader is invited to actively work on the material.

W.2 Quantizing a Field - A Toy Example

We present a simple toy example in order to clarify the basic ideas of field quantization. The system consists of two conducting plates at $z = 0$ and $z = L$ which are infinitely extended in $x - y$-direction. Between the plates, we assume an electrical field in the form of standing waves in z-direction which are linearly polarized in x-direction. Since the plates are conducting, the electrical field vanishes at $z = 0$ and $z = L$.

We start from the source-free Maxwell equations

$$\begin{array}{c} \nabla\cdot\mathbf{E}(\mathbf{r},t) = 0 \;;\; \nabla\cdot\boldsymbol{B}(\mathbf{r},t) = 0 \\ \nabla\times\mathbf{E}(\mathbf{r},t) = -\frac{\partial}{\partial t}\mathbf{B}(\mathbf{r},t) \;;\; \nabla\times\mathbf{B}(\mathbf{r},t) = \frac{1}{c^2}\frac{\partial}{\partial t}\mathbf{E}(\mathbf{r},t)\,. \end{array} \tag{W.1}$$

W.2.1 The Classical Case

According to our assumptions, the electrical field is given by

$$\mathbf{E}(\mathbf{r},t) = (E_x(z,t), 0, 0) \text{ with } E_x(0,t) = E_x(L,t) = 0 \tag{W.2}$$

where $E_x(z,t)$ may be written as

$$E_{xl}(z,t) = C_l \cdot q_l(t)\sin k_l z \;;\; k_l = \frac{\pi}{L}l \text{ with } l = 1, 2, \ldots. \tag{W.3}$$

C_l is a normalization factor, $q_l(t)$ the amplitude of the field (yet to be determined). Due to the last two Maxwell equations, for the magnetic field holds $\mathbf{B}(\mathbf{r},t) = \left(0, B_y(z,t), 0\right)$ with

$$B_{yl}(z,t) = C_l \cdot \frac{1}{c^2 k_l}\dot{q}_l(t)\cos k_l z. \tag{W.4}$$

Of course, the general solution for e.g. the electrical field is the linear superposition of all partial solutions (W.3), i.e., the sum over all l. However, in order to hold the discussion transparent, we consider for the moment only *one* partial solution for a definite l.

The general expression for the electromagnetic energy H is given by $H = \frac{\varepsilon_0}{2}\int d^3x\,\left(\mathbf{E}^2 + c^2\mathbf{B}\right)$. In our simple case, this reads

$$H_l = \frac{\varepsilon_0}{2}\int_0^L dz\left(E_{xl}^2 + c^2 B_{yl}^2\right). \tag{W.5}$$

Evaluation of the integral leads with $\omega_l^2 = k_l^2 c^2$ to

$$H_l = \frac{1}{2}\left[\dot{q}_l^2(t) + \omega_l^2 q_l^2(t)\right] \tag{W.6}$$

where we have fixed the normalization constant by $C_l = \sqrt{\frac{2\omega_l^2}{\varepsilon_0 L}}$. This expression is formally identically equal to the energy of a harmonic oscillator (of a fictive mass $m = 1$), if we identify $\dot{q}_l$ with p_l.

W.2.2 Quantization

Quantization means, that we interpret the classical quantities p_l and q_l as operators[67] $\hat{p}_l$ and $\hat{q}_l$ which obey the commutation rule $\left[\hat{q}_l, \hat{p}_l\right] = i\hbar$. In this way, we arrive at the quantized version of our toy system

$$\hat{H}_l = \frac{1}{2}\left[\hat{p}_l^2(t) + \omega_l^2 \hat{q}_l^2(t)\right] \;;\; \left[\hat{q}_l, \hat{p}_l\right] = i\hbar. \tag{W.7}$$

Of course, the fields become operators, too (hence called field operators):

$$\hat{E}_{xl}(z,t) = C_l \cdot \hat{q}_l(t) \sin k_l z \;;\; \hat{B}_{yl}(z,t) = \frac{1}{c^2 k_l} C_l \cdot \hat{p}_l(t) \cos k_l z. \tag{W.8}$$

Note that with $\hat{p}_l$ and $\hat{q}_l$ also the fields are hermitian operators and as such observables. Let us point out, that now the field operators do not commute:

$$\begin{aligned} &\left[\hat{E}_{xl}(z,t), \hat{B}_{yl}(z,t)\right] = \hat{E}_{xl}(z,t)\hat{B}_{yl}(z,t) - \hat{B}_{yl}(z,t)\hat{E}_{xl}(z,t) = \\ &= \tfrac{C_l^2}{c^2 k_l} \sin k_l z \cos k_l z \left[\hat{q}_l(t) \cdot \hat{p}_l(t) - \hat{p}_l(t) \cdot \hat{q}_l(t)\right] = i\hbar \tfrac{C_l^2}{c^2 k_l} \sin k_l z \cos k_l z. \end{aligned} \tag{W.9}$$

W.2.3 Creation and Annihilation Operators, Hamiltonian

We now can apply the whole machinery of the harmonic oscillator as developed in Chap. 18, Vol. 2. We define a creation (raising) operator $a^\dagger(k_l)$ and an annihilation (lowering) operator $a(k_l)$ by (as usual, these ladder operators are written without hat)

$$\begin{aligned} a^\dagger(k_l) &= \tfrac{1}{\sqrt{2\hbar\omega_l}}\left(\omega_l \hat{q}_l - i\hat{p}_l\right) \;;\; a(k_l) = \tfrac{1}{\sqrt{2\hbar\omega_l}}\left(\omega_l \hat{q}_l + i\hat{p}_l\right) \\ \hat{q}_l &= \sqrt{\tfrac{\hbar}{2\omega_l}}\left(a^\dagger(k_l) + a(k_l)\right) \;;\; \hat{p}_l = i\sqrt{\tfrac{\hbar\omega_l}{2}}\left(a^\dagger(k_l) - a(k_l)\right) \end{aligned} \tag{W.10}$$

and the commutation relations read

$$\left[a(k_l), a^\dagger(k_l)\right] = 1 \;;\; \left[\hat{q}_l, \hat{p}_l\right] = i\hbar. \tag{W.11}$$

The field operators are given by

$$\hat{E}_{xl}(z,t) = C_l\sqrt{\frac{\hbar}{2\omega_l}}\left[a^\dagger(k_l) + a(k_l)\right]\sin k_l z \;;\; \hat{B}_{yl}(z,t) = C_l \frac{i}{c^2 k_l}\sqrt{\frac{\hbar\omega_l}{2}}\left[a^\dagger(k_l) - a(k_l)\right]\cos k_l z. \tag{W.12}$$

With (W.7) and (W.10) we obtain the Hamiltonian $\hat{H}_l$ in the form

[67] In this section, we denote operators by a hat.
The quantization procedure is canonical quantization; cf. Vol. 1, App. T.3.

$$\hat{H}_l = \hbar\omega_l \frac{a^\dagger(k_l)a(k_l) + a(k_l)a^\dagger(k_l)}{2}. \quad \text{(W.13)}$$

One can show that the product $aa^\dagger$ is time-independent and therefore also $\hat{H}_l$ (see exercises). Hence, we can suppress the (uninteresting) time dependence and keep the notation $a(k_l)$.

With the help of the commutation relation (W.11), i,e., $a(k_l)a^\dagger(k_l) = 1 - a^\dagger(k_l)a(k_l)$, we rewrite $\hat{H}_l$ in (W.13) and obtain

$$\hat{H}_l = \hbar\omega_l \left[a^\dagger(k_l)a(k_l) + \frac{1}{2} \right] \quad \text{(W.14)}$$

which expression formally equals the energy of the harmonic oscillator. In addition, we can also define a number operator $\hat{N}_l = a^\dagger(k_l)a(k_l)$ with eigenstates $|n_l\rangle$; the eigenwert equation reads $N_l\,|n_l\rangle = n_l\,|n_l\rangle$. The states $|n_l\rangle$ are orthogonal, $\langle n_l|\,n_j\rangle = \delta_{n_l n_j}$.

Due to

$$\hat{H}_l\,|n_l\rangle = \hbar\omega_l \left(\hat{N}_l + 1/2 \right) |n_l\rangle = \hbar\omega_l\,(n_l + 1/2)\,|n_l\rangle = E_{l,n_l}\,|n_l\rangle \quad \text{(W.15)}$$

we can write for the Hamilton function and for the energy[68]

$$\hat{H}_l = \hbar\omega_l \left(\hat{N}_l + \frac{1}{2} \right) \; ; \; E_{l,n_l} = \hbar\omega_l \left(n_l + \frac{1}{2} \right). \quad \text{(W.16)}$$

The ground state is $|0\rangle$ (the vacuum), and the ladder operators rise and lower the states, $a^\dagger(k_l)\,|n_l\rangle = \sqrt{n_l + 1}\,|n_l + 1\rangle$ and $a(k_l)\,|n_l\rangle = \sqrt{n_l}\,|n_l - 1\rangle$. A state $|n_l\rangle$ can be produced out of the vacuum by repeated application of $a^\dagger(k_l)$:

$$a^\dagger(k_l)\,|0\rangle = \sqrt{1}\,|1\rangle \;\; ; \; a^\dagger(k_l)\,|1\rangle = \left[a^\dagger(k_l)\right]^2 |0\rangle = \sqrt{2}\,|1\rangle \;\; ; \ldots \left[a^\dagger(k_l)\right]^m |0\rangle = \sqrt{m!}\,|m\rangle \quad \text{(W.17)}$$

or

$$|n_l\rangle = \frac{1}{\sqrt{n_l!}} \left[a^\dagger(k_l)\right]^{n_l} |0\rangle\,. \quad \text{(W.18)}$$

Remind that the application of $a(k_l)$ to the vacuum vanishes, $a(k_l)\,|0\rangle = 0$.

Though the formalism is the same for the harmonic oscillator[69] and our toy example, the interpretation differs. In case of the harmonic oscillator, $E_m = \hbar\omega\left(m + \frac{1}{2}\right)$ means that the oscillator is in the mth level, but we now interpret $E_{l,n_l} = \hbar\omega_l\left(n_l + \frac{1}{2}\right)$ as an indication that the mode[70] is occupied by n_l 'particles' of energy $\hbar\omega_l$. These

---

[68] Do not confuse energy E_n and electric field E_x.

[69] We remark that it is not at all self-evident that we can use here the formalismus of the harmonic oscillator. It is rather one of those serendipities in physics.

[70] Here, a mode of the field is determined by its energy $\hbar\omega_l$.

'particles' are named *photons*. Note that 'photon' means just the basic unit of the energy of the electromagnetic field and not some sort of point particle running through the space. Indeed, the energy of a mode is delocalized and property of the mode.

Thus, e.g. the state $|n_l\rangle$ contains n_l photons of energy $\hbar\omega_l$ - in other words, $|n_l\rangle$ is the state in which the mode is occupied by n_l photons.

W.2.4 Generalization

In our toy example, we have assumed two plates between which there are standing waves in z-direction which are linearly polarized in x-direction. We now generalize the example. We assume that we have not only two plates but a cube of length L. Thus, the allowed wave vectors are given by

$$\mathbf{k} = \frac{\pi}{L}\mathbf{n} \;\; ; \; n_x, n_y, n_z \in \mathbb{Z}. \tag{W.19}$$

In addition, we allow for arbitrary plane waves (i.e., fields with three non-vanishing components) and for arbitrary polarizations (not only in x-direction). Thus, a mode is determined by the numbers $(\mathbf{k}, r)$ where $r = 1, 2, 3$ indicates the polarization direction (e.g., in x-, y- and z-direction). Thus, we write the creation operator (formerly $a(k_l)$) as $a_r(\mathbf{k})$. In addition, instead of the partial Hamilton function H_l (W.14) for one mode, we consider all modes $(\mathbf{k}, r)$. Thus, with a similar calculation as above we arrive at the Hamilton operator $\hat{H}$:

$$\hat{H} = \sum_{\mathbf{k},r} \hbar\omega_{\mathbf{k}} \left[a_r^\dagger(\mathbf{k})\, a_r(\mathbf{k}) + \frac{1}{2} \right]. \tag{W.20}$$

Here, the summation runs over all allowed $\mathbf{k}$-values and all polarization directions r, and $\omega_{\mathbf{k}} = c\,|\mathbf{k}|$. Each mode behaves like a independent harmonic oscillator and can accept energy in an integer number of portions (quanta) of size $\hbar\omega_{\mathbf{k}}$. The commutation relation (W.11) takes on the form

$$\left[a_r^\dagger(\mathbf{k}), a_r(\mathbf{k}) \right] = \delta_{\mathbf{k}\mathbf{k}'}\delta_{rr'}. \tag{W.21}$$

The number operator is defined by

$$N_{\mathbf{k}r} = a_r^\dagger(\mathbf{k})\, a_r(\mathbf{k}) \tag{W.22}$$

measuring the number of modes with quantum numbers $\mathbf{k}r$. We denote by $|n_{\mathbf{k}r}\rangle$ the eigenvectors of $N_{\mathbf{k}r}$, and the eigenvalue equation reads[71]

[71] A more detailed notation would be $|\mathbf{k}r, n_{\mathbf{k}r}\rangle$, i.e., a state with quantum numbers $\mathbf{k}r$ and the occupation number $n_{\mathbf{k}r}$. Instead, we prefer the shorter and equivalent notation $|n_{\mathbf{k}r}\rangle$.

$$N_{\mathbf{k}r} |n_{\mathbf{k}r}\rangle = n_{\mathbf{k}r} |n_{\mathbf{k}r}\rangle \tag{W.23}$$

where $n_{\mathbf{k}r}$ is the occupation number of the mode indicated by $(\mathbf{k}r)$. As in the simpler version of our example, we can adopt the interpretation that we have $n_{\mathbf{k}r}$ photons with wave vector $\mathbf{k}$ and polarization r - or: the mode $(\mathbf{k}, r)$ is occupied by $n_{\mathbf{k}r}$ photons. Application of the operators $a_r^\dagger(\mathbf{k})$ and $a_r(\mathbf{k})$ increases and lowers this number by one, and we can represent the states by repeated application of the creation operator $a_r^\dagger(\mathbf{k})$:

$$|n_{\mathbf{k}r}\rangle = \frac{1}{\sqrt{n_{\mathbf{k}r}!}} \left[a_r^\dagger(\mathbf{k})\right]^{n_{\mathbf{k}r}} |0\rangle \tag{W.24}$$

and

$$a_r(\mathbf{k}) |0\rangle = 0. \tag{W.25}$$

A general state which comprises all modes may be written $|n_1, n_2, n_3, \ldots\rangle$, i.e., the first mode is occupied by n_1 photons, the second by n_2 and so on. In compact form, this may be written as

$$|n_1, n_2, n_3, \ldots\rangle = \prod_{\mathbf{k},r} \frac{1}{\sqrt{n_{\mathbf{k}r}!}} \left[a_r^\dagger(\mathbf{k})\right]^{n_{\mathbf{k}r}} |0\rangle . \tag{W.26}$$

In other words, we do not label the (identical) photons by assigning each of them an individual quantum state.[72] Instead, we count how many photons are occupying a mode.

This notation and its interpretation is known as *occupation number representation*. As seen from (W.26), it enables us to go almost without state vectors except the vacuum state $|0\rangle$. One says that the operators $a_r^\dagger(\mathbf{k})$ and $a_r(\mathbf{k})$ create and annihilate a particle (photon) with quantum numbers $(\mathbf{k}, r)$. The states $|n\rangle$ are called *number states* or *Fock states* since they live in Fock space.

W.2.4.1 Summary of the Quantization Approach

Let us resume the approach. We started with the classical field equations and represented the solution in terms of plane waves whose amplitudes were essentially q and $p = \dot{q}$. Thus, we could identify by mere inspection q and p as canonical variables and transform them into non-commuting operators. This means that also the fields are transformed into field operators. The energy of the system is a function of q and p and formally identical to that of the harmonic oscillator. Exploiting the formal analogy, we can use the formalism of the harmonic operator to construct ladder, number and energy operators though the two systems (toy example and harmonic oscillator) are physically completely different. In contrast to the harmonic oscillator, we adopt

[72] By the way, this would not make sense since photons are indistinguishable; see Vol. 2 Chap. 23 'Identical Particles'.

for our toy system the occupation number representation which counts how many photons are occupying a single mode.

In principle, this is the method to quantize all fields. However, the way chosen here is tailored to our toy system and we have to find a general approach by which we can identify the canonical variables. In possession of these terms, we can formulate creation and annihilation operators and all other quantities of interest. As we will see below, this tool is the Lagrange–Hamilton formalism.

W.2.4.2 A Big Problem

So far, everything is coherent and fine - except for a 'small' problem, which in fact is an infinite one. The energy associated with the Hamiltonian (W.20) is given by

$$E = \sum_{\mathbf{k},r} \hbar\omega_{\mathbf{k}} \left(n_{\mathbf{k}r} + \frac{1}{2} \right). \tag{W.27}$$

We can split and write this as

$$E = \sum_{\mathbf{k},r} \hbar\omega_{\mathbf{k}}\, n_{\mathbf{k}r} + \sum_{\mathbf{k},r} \frac{\hbar\omega_{\mathbf{k}}}{2}. \tag{W.28}$$

Due to the second summand, i.e., the sum over the zero point energies, the energy E is *always infinite*, even if all occupation numbers $n_{\mathbf{k}r}$ vanish. Let us point out in advance that this is not a speciality of our toy system. In fact, all fields to be quantized in the following (Klein–Gordon, Dirac, photons) display the same problem of infinite zero point or vacuum energy.

Of course, this is a serious problem. Can a formalism be credible in which one always has to take account of an infinitely large number - or formulated differently: in which one has to subtract an infinite number to get finite results? In general, physicists are rather nonchalant with their mathematical methods, and sometimes allow for some sloppiness for the sake of argument. But this divergence problem literally has a different order of magnitude.

We postpone the discussion and address the subject again in section 'Operator ordering'.

W.2.5 Exercises and Solutions

1. Show that (W.3) and (W.4) satisfy the Maxwell equations (W.1).
   Solution: We have for the divergence terms

$$\nabla\cdot\mathbf{E}\,(\mathbf{r},t) = \partial_x E_x\,(z,t) = 0 \;\; ; \;\; \nabla\cdot\mathbf{B}\,(\mathbf{r},t) = \partial_y B_y\,(z,t) = 0. \tag{W.29}$$

The rotation terms are given by $\nabla \times \mathbf{E}(\mathbf{r},t) = (0, \partial_z E_x, 0)$ and $\nabla \times \mathbf{B}(\mathbf{r},t) = \left(-\partial_z B_y, 0, 0\right)$. It follows

$$\partial_z E_x(z,t) = -\frac{\partial}{\partial t} B_y(z,t) \to k_m C \cdot q(t) \cos k_m z = -C \cdot \frac{1}{c^2 k_m} \ddot{q}(t) \cos k_m z \tag{W.30}$$

and

$$-\partial_z B_y(z,t) = \frac{1}{c^2} \frac{\partial}{\partial t} E_x(z,t) \to C \cdot \frac{1}{c^2} \dot{q}(t) \sin k_m z = \frac{1}{c^2} C \cdot \dot{q}(t) \sin k_m z. \tag{W.31}$$

The last equation is always satisfied, and the first for $\ddot{q}(t) = -c^2 k_m^2 q(t)$.

2. Prove (W.6)., i.e.,

$$H = \frac{1}{2}\left[\dot{q}^2(t) + \omega_m^2 q^2(t)\right]. \tag{W.32}$$

Solution: We have

$$H = \frac{\varepsilon_0}{2} \int_0^L \mathrm{d}z \left[E_x^2 + c^2 B_y^2\right] = \frac{\varepsilon_0}{2} \int_0^L \mathrm{d}z \left[C^2 q^2(t) \sin^2 k_m z + c^2 C^2 \frac{1}{c^4 k_m^2} \dot{q}^2(t) \cos^2 k_m z\right]. \tag{W.33}$$

This yields

$$H = \frac{\varepsilon_0}{2} C^2 \left[q^2(t) \int_0^L \mathrm{d}z \sin^2 k_m z + \frac{1}{c^2 k_m^2} \dot{q}^2(t) \int_0^L \mathrm{d}z \cos^2 k_m z\right]. \tag{W.34}$$

With

$$\int_0^L \mathrm{d}z \sin^2 k_m z = \left\{\frac{z}{2} - \frac{\sin 2k_m z}{4k_m}\right\}_0^L = \frac{L}{2} - \frac{\sin 2\frac{\pi}{L} mL}{4k_m} = \frac{L}{2} \tag{W.35}$$

and

$$\int_0^L \mathrm{d}z \cos^2 k_m z = \int_0^L \mathrm{d}z \left(1 - \sin^2 k_m z\right) = \frac{L}{2} \tag{W.36}$$

follows with $\omega_m^2 = c^2 k_m^2$

$$H = \frac{\varepsilon_0}{2} C^2 \frac{L}{2} \left[q^2(t) + \frac{1}{c^2 k_m^2} \dot{q}^2(t)\right] = \frac{\varepsilon_0 L}{2\omega_m^2} C^2 \frac{1}{2} \left[\dot{q}^2(t) + \omega_m^2 q^2(t)\right]. \tag{W.37}$$

The choice $C^2 = \frac{2\omega_m^2}{\varepsilon_0 L}$ brings the desired result.

3. Show (W.13).

Solution: We have with (W.7) and (W.10)

$$\begin{aligned}\hat{H}_l &= \tfrac{1}{2}\left[-\tfrac{\hbar\omega_l}{2}\left(a^\dagger(k_l)-a(k_l)\right)\left(a^\dagger(k_l)-a(k_l)\right)+\omega_l^2\tfrac{\hbar}{2\omega_l}\left(a^\dagger(k_l)+a(k_l)\right)\left(a^\dagger(k_l)+a(k_l)\right)\right]=\\ &= \tfrac{\hbar\omega_l}{4}\left[\begin{array}{c}-a^\dagger(k_l)a^\dagger(k_l)+a^\dagger(k_l)a(k_l)+a(k_l)a^\dagger(k_l)-a(k_l)a(k_l)+\\ +\left(a^\dagger(k_l)a^\dagger(k_l)+a(k_l)a^\dagger(k_l)+a(k_l)a^\dagger(k_l)+a(k_l)a(k_l)\right)\end{array}\right]=\\ &= \tfrac{\hbar\omega_l}{2}\left[a^\dagger(k_l)a(k_l)+a(k_l)a^\dagger(k_l)\right].\end{aligned} \tag{W.38}$$

4. Show that $a(k_l)a^\dagger(k_l)$ is time-independent.
Solution: Let us write for the moment $a(k_l, t)$. We invoke the definition of the time derivative in the Heisenberg picture as given by $\mathrm{d}A/\mathrm{d}t = i/\hbar\,[H, Aa]$. In our case this leads to

$$\begin{aligned}\frac{\mathrm{d}a(k_l,t)}{\mathrm{d}t} &= \frac{i}{\hbar}\left[\hat{H}_l, a(k_l,t)\right] = i\omega_l\left[a^\dagger(k_l,t)a(k_l,t), a(k_l,t)\right]\\ &= -i\omega_l a(k_l,t)\ ;\ \frac{\mathrm{d}a^\dagger(k_l,t)}{\mathrm{d}t} = i\omega_l a^\dagger(k_l,t)\end{aligned} \tag{W.39}$$

with the solutions

$$a(k_l,t) = a(k_l,0)\cdot e^{-i\omega_l t}\ ;\ a^\dagger(k_l,t) = a^\dagger(k_l,0)\cdot e^{i\omega_l t} \tag{W.40}$$

and it follows

$$a(k_l,t)a^\dagger(k_l,t) = a(k_l,0)a^\dagger(k_l,0). \tag{W.41}$$

Thus, we see clearly that the product $aa^\dagger$ is time-independent and therefore also $\hat{H}_l$. Hence, we switch back to the notation $a(k_l)$, suppressing the (now uninteresting) time dependence.

W.3 Quantization of Free Fields, Introduction

In the last section, we used an approach for the quantization which was tailored to the system under consideration. We discuss now the general method (in fact, *the* general method) which is based on the Lagrange–Hamilton formalism.[73] It answers the relevant questions as, for instant, how to find those variables which are transformed into non-commuting operators, how to find the energy density (Hamiltonian density) and so on.

The Lagrangian contains the complete information about the physical system. It enables us to derive (1) the equations of motions, (2) the conjugated momentum, (3) the Hamiltonian. Moreover, it forms the basis of the canonical quantization by which the field operator and its conjugated momentum are subject to certain commutation relations.

In addition, following our approach for the toy system, we have a further point in the quantizing procedure, namely to express the field operators in terms of anni-

[73] In Vol. 1, App. T.3, there is a short outline of the basics of this formalism.

hilation and creation operators. To this end, we construct the free solutions and replace the expansion coefficients by operators. The final step is the formulation of the commutation relation for these annihilation and creation operators.

We summarize the *canonical quantization* procedure which we will apply in case of the Klein–Gordon field. Assume that we have a physical (classical, non-quantum) system with an appropriate Lagrangian $\mathcal{L}$ for several fields φ_r, $r = 1, 2, \ldots$. Then we have to carry out the following steps:

1. We calculate the conjugated momentum fields π_r and formulate the corresponding Hamiltonian $\mathcal{H}$ density as a function of the fields φ_r and the momentum densities π_r.

2. We consider φ_r and π_r as operators obeying certain commutation relations.

3. On the basis of free solutions of the system, we formulate φ_r and π_r in terms of annihilation and creation operators and deduce the commutation relations for these two types of operators.

However, this approach is, in a certain sense, idealized and/or of limited value. The reason is that it is based on the knowledge of $\mathcal{L}$ (or $\mathcal{H}$) of the classical system to be quantized. But there are systems with no classical $\mathcal{L}$ as for instance the spinor field of the Dirac equation. Thus, in this case, the approach can not be applied directly, and we have to look for the right Lagrangian, not to mention the appropriate commutation relations.

In the following, we will first consider the Klein–Gordon field as example for the canonical quantization. Then we will show for the Dirac system how to proceed if there is no classical Lagrangian. Finally, we compile some results for the free photon field.

Starting with this section, we will use those physical units in which $c = 1$ and $\hbar = 1$ (see Appendix B, Vol. 1, 'Natural units'). This is common practice in quantum field theory and very functional. As a consequence, several quantities are now directly equivalent. We have for instance $\mathbf{p} = \mathbf{k}$ or $E_\mathbf{k} = \omega_\mathbf{k}$, the dispersion relation may be written $E_\mathbf{k} = \sqrt{\mathbf{k}^2 + m^2}$, and for the 4-vector p holds $p^0 = k^0 = E_\mathbf{k} = \omega_\mathbf{k} = E_\mathbf{p} = \omega_\mathbf{p}$.

W.4 Quantization of Free Fields, Klein–Gordon

We now consider the real Klein–Gordon field and its quantization. This field describes electrical neutral mesons with spin zero which are equal to their antiparticles.[74] Apart from its physical importance, the Klein–Gordon field is a nice example for canonical quantization.

[74] Mesons which differ from their antiparticles (e.g. the K_0 meson with its antiparticle $\bar{K}_0$) are described by a complex Klein–Gordon field.

W.4.1 Lagrangian, Conjugated Momentum, Poisson Brackets, Hamiltonian

The Lagrangian density $\mathcal{L}$ of the real Klein–Gordon field is given by[75]

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \phi\right)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2. \tag{W.42}$$

The Euler–Lagrange equation reads

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi\right)} \right) = 0. \tag{W.43}$$

With

$$\frac{\partial \mathcal{L}}{\partial \phi} = -m^2 \phi \ ; \ \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi\right)} = \partial^\mu \phi \tag{W.44}$$

follows the Klein–Gordon equation

$$\partial_\mu \partial^\mu \phi + m^2 \phi = 0. \tag{W.45}$$

The conjugated momentum density π is given by

$$\pi = \frac{\delta \mathcal{L}}{\delta \left(\partial_0 \phi\right)} = \partial_0 \phi \tag{W.46}$$

and the Hamiltonian density follows as

$$\mathcal{H} = \frac{1}{2}\left[\dot{\phi}^2 + (\nabla \phi)^2 + m^2 \phi^2\right]. \tag{W.47}$$

W.4.2 Canonical Quantization

The Poisson brackets for the Klein–Gordon field are given by[76] (see Appendix T, Vol. 1)

---

[75] Remarks concerning the notation: Instead of $\left(\partial_\mu \phi\right)(\partial^\mu \phi)$, some textbooks write $\left(\partial_\mu \phi\right)^2$. One finds also $\partial_\mu \phi \partial^\mu \phi$, whereby this is not meant in the sense of the product rule, i.e., $\left(\partial_\mu \phi\right)(\partial^\mu \phi) + \phi \partial_\mu \partial^\mu \phi$. In any case, the following expression is meant: $(\partial_0 \phi)\left(\partial^0 \phi\right) - (\partial_k \phi)(\partial_k \phi) = (\partial_0 \phi)^2 - (\partial_k \phi)^2 = \dot{\phi}^2 - (\nabla \phi)^2$. In these terms, (W.42) is written $\mathcal{L} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla \phi)^2 - \frac{1}{2} m^2 \phi^2$.

[76] In order to emphasize the dimension of the argument, we write occasionally $\delta^{(3)}(\mathbf{x})$ instead of $\delta(\mathbf{x})$; analogously for other dimensions.

$$\begin{aligned} \left\{\phi(t,\mathbf{x}), \pi\left(t,\mathbf{x}'\right)\right\}_{PB} &= \delta^{(3)}\left(\mathbf{x}-\mathbf{x}'\right) \\ \left\{\phi(t,\mathbf{x}), \phi\left(t,\mathbf{x}'\right)\right\}_{PB} &= 0 \\ \left\{\pi(t,\mathbf{x}), \pi\left(t,\mathbf{x}'\right)\right\}_{PB} &= 0. \end{aligned} \tag{W.48}$$

The canonical quantization of Klein–Gordon field means that we consider the field variables $\phi(x)$ and $\pi(x)$ as operators obeying the commutation relations[77]

$$\begin{aligned} \left[\phi(t,\mathbf{x}), \dot{\phi}\left(t,\mathbf{x}'\right)\right] &= i\delta^{(3)}\left(\mathbf{x}-\mathbf{x}'\right) \\ \left[\phi(t,\mathbf{x}), \phi\left(t,\mathbf{x}'\right)\right] &= \left[\dot{\phi}(t,\mathbf{x}), \dot{\phi}\left(t,\mathbf{x}'\right)\right] = 0. \end{aligned} \tag{W.49}$$

Note the additional factor i in the transition from the Poisson brackets to the commutation rules. We now expand $\phi(x)$ in terms of plane wave solutions of the Klein–Gordon equation (see Appendix U, Vol. 1). In the case of a finite volume V, they read

$$\phi(x) = \sum_{\mathbf{p}} \frac{1}{\sqrt{2VE_{\mathbf{p}}}} \left(a(\mathbf{p})\, e^{-ipx} + a^{\dagger}(\mathbf{p})\, e^{ipx}\right) \tag{W.50}$$

where the sum runs over all allowed discrete values of $\mathbf{p}$. The continuous solution reads[78]

$$\phi(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 p}{\sqrt{2E_{\mathbf{p}}}} \left(a(\mathbf{p})\, e^{-ipx} + a^{\dagger}(\mathbf{p})\, e^{ipx}\right). \tag{W.51}$$

Energy and momentum are related by the relativistic dispersion relation

$$E_{\mathbf{p}} = p^0 = \sqrt{\mathbf{p}^2 + m^2}. \tag{W.52}$$

Inverting the continuous solution (W.51) gives for the operators $a(\mathbf{p})$ and $a^{\dagger}(\mathbf{p})$

$$\begin{aligned} a(\mathbf{p}) &= \frac{1}{\sqrt{(2\pi)^3 2E_{\mathbf{p}}}} \int \mathrm{d}^3x\, e^{-i\mathbf{px}} \left(E_{\mathbf{p}}\phi(0,\mathbf{x}) + i\dot{\phi}(0,\mathbf{x})\right) \\ a^{\dagger}(\mathbf{p}) &= \frac{1}{\sqrt{(2\pi)^3 2E_{\mathbf{p}}}} \int \mathrm{d}^3x\, e^{i\mathbf{px}} \left(E_{\mathbf{p}}\phi(0,\mathbf{x}) - i\dot{\phi}(0,\mathbf{x})\right). \end{aligned} \tag{W.53}$$

From (W.49) and (W.53) follow the commutation relations for $a(\mathbf{p})$ and $a^{\dagger}(\mathbf{p})$:

$$\left[a(\mathbf{p}), a^{\dagger}\left(\mathbf{p}'\right)\right] = \delta\left(\mathbf{p},\mathbf{p}'\right)\ ;\ \left[a(\mathbf{p}), a\left(\mathbf{p}'\right)\right] = \left[a^{\dagger}(\mathbf{p}), a^{\dagger}\left(\mathbf{p}'\right)\right] = 0 \tag{W.54}$$

which are again (formally) identical with the commutation relations of a harmonic oscillator[79]In other words: we can now follow the considerations from the toy

---

[77]Here the fields are considered at the same time. For different times, the fields commute.

[78]As in Appendix U, Vol. 1, we use the same symbol $a(\mathbf{p})$ in the discrete and the continuous case. Strictly speaking one would have to make a distinction e.g. by different names. But our approach is quite functional and the likelihood of confusion is very low.

[79]$\delta(a,b)$ is the generalized Kronecker symbol introduced in Vol. 1, Chap. 12:

Table W.1 Table of simplest states of the Klein–Gordon field

State		Energy
Vacuum	$\|0\rangle$	0
One particle	$a^{\dagger}(\mathbf{p})\|0\rangle$	$E_{\mathbf{p}}$
Two different particles	$a^{\dagger}(\mathbf{p}_1)\, a^{\dagger}(\mathbf{p}_2)\|0\rangle$	$E_{\mathbf{p}_1}+E_{\mathbf{p}_2}$
Two identical particles	$a^{\dagger}(\mathbf{p})\, a^{\dagger}(\mathbf{p})\|0\rangle$	$2E_{\mathbf{p}}$

example, i.e., we can adopt the formalism of the harmonic oscillator and the occupation number representation.

One can regard the operators

$$N_{\mathbf{p}} = a^{\dagger}(\mathbf{p})\, a(\mathbf{p}) \tag{W.56}$$

as number operators with eigenvalues 0, 1, 2, ... The operators $a(\mathbf{p})$ and $a^{\dagger}(\mathbf{p})$ delete and create particles with momentum p.[80] For instance, $a^{\dagger}(\mathbf{p})|0\rangle$ creates a particle with momentum p, and $a^{\dagger}(\mathbf{p}_1)\, a^{\dagger}(\mathbf{p}_2)|0\rangle$ two particles with momenta p_1 and p_2. Some normalized particle states with their energy eigenvalues are given in Table W.1.

Finally, we look for the Hamiltonian. We have with (W.47)

$$H = \int \mathrm{d}^3x\, \mathcal{H}(x) = \int \mathrm{d}^3x\, \frac{1}{2}\left[\dot{\phi}^2(x) + (\nabla\phi(x))^2 + m^2\phi^2(x)\right]. \tag{W.57}$$

With (W.50) or with (W.51), this leads to

$$H = \frac{1}{2}\sum_{\mathbf{p}} E_{\mathbf{p}}\left(a^{\dagger}(\mathbf{p})\, a(\mathbf{p}) + a(\mathbf{p})\, a^{\dagger}(\mathbf{p})\right) \tag{W.58}$$

or

$$H = \frac{1}{2}\int \mathrm{d}^3p\, E_{\mathbf{p}}\left[a^{\dagger}(\mathbf{p})\, a(\mathbf{p}) + a(\mathbf{p})\, a^{\dagger}(\mathbf{p})\right]. \tag{W.59}$$

Two remarks:

(1) Note that there are only *positive* energies because the number operator $a^{\dagger}(\mathbf{p})\, a(\mathbf{p})$ has only positive eigenvalues (plus zero). In other words, we got rid of the problem with negative energies.

(2) Instead, we face another problem, known from the toy example: the energy is infinite always. By means of (W.54) we can write for example

$$\delta(a,b) = \begin{cases} \delta_{ab} & \text{for } a, b \text{ discrete} \\ \delta(a-b) & \text{for } a, b \text{ continuous} \end{cases} \tag{W.55}$$

[80] Note that due to $p_0^2 = \mathbf{p}^2 + m^2$, it does not matter if we give the 4-momentum p or the 3-momentum $\mathbf{p}$.

$$H = \sum_{\mathbf{p}} \frac{1}{2} E_{\mathbf{p}} \left(a^{\dagger}(\mathbf{p})\, a(\mathbf{p}) + a(\mathbf{p})\, a^{\dagger}(\mathbf{p}) \right) = \sum_{\mathbf{p}} E_{\mathbf{p}} \left(a^{\dagger}(\mathbf{p})\, a(\mathbf{p}) + \frac{1}{2} \right) \quad \text{(W.60)}$$

This means that even if there is no particle at all (i.e., vacuum), there is a energy of $\frac{1}{2}\sum_{\mathbf{p}} E_{\mathbf{p}}$ (which we have omitted in the last table). We will find this issue also in the case of the Dirac field and the photon field.

Thus, there is a welcome property (energies are only positive) and an unwelcome property (infinite energies). We will discuss this problem below in the section 'Operator ordering'.

W.4.3 Exercises and Solutions

1. Show

$$\frac{\partial}{\partial\left(\partial_{\mu}\phi\right)} \left(\partial_{\alpha}\phi\right)\left(\partial^{\alpha}\phi\right) = 2\left(\partial^{\mu}\phi\right). \quad \text{(W.61)}$$

Solution: We have

$$\begin{aligned} &\frac{\partial}{\partial(\partial_{\mu}\phi)} \left(\partial_{\alpha}\phi\right)\left(\partial^{\alpha}\phi\right) = \delta_{\mu\alpha}\left(\partial^{\alpha}\phi\right) + \left(\partial_{\alpha}\phi\right) \frac{\partial}{\partial(\partial_{\mu}\phi)} \left(g^{\alpha\nu}\partial_{\nu}\phi\right) = \\ &= \left(\partial^{\mu}\phi\right) + \left(\partial_{\alpha}\phi\right) g^{\alpha\nu}\delta_{\mu\nu}\phi = \left(\partial^{\mu}\phi\right) + \left(\partial^{\nu}\phi\right)\delta_{\mu\nu}\phi = \left(\partial^{\mu}\phi\right) + \left(\partial^{\mu}\phi\right) = 2\left(\partial^{\mu}\phi\right). \end{aligned} \quad \text{(W.62)}$$

2. Prove (W.47).
Solution:

$$\begin{aligned} \mathcal{H} &= \pi\partial_0\phi - \mathcal{L} = \left(\partial_0\phi\right)^2 - \tfrac{1}{2}\left(\partial_{\mu}\phi\right)\left(\partial^{\mu}\phi\right) + \tfrac{1}{2}m^2\phi^2 = \\ &= \left(\partial_0\phi\right)^2 - \tfrac{1}{2}\left(\partial_0\phi\right)\left(\partial^0\phi\right) - \tfrac{1}{2}\left(\partial_k\phi\right)\left(\partial^k\phi\right) + \tfrac{1}{2}m^2\phi^2 = \\ &= \tfrac{1}{2}\left(\partial_0\phi\right)\left(\partial^0\phi\right) - \tfrac{1}{2}\left(\partial_k\phi\right)\left(\partial^k\phi\right) + \tfrac{1}{2}m^2\phi^2 = \tfrac{1}{2}\left[\left(\partial_0\phi\right)^2 + \left(\partial_k\phi\right)\left(\partial^k\phi\right) + m^2\phi^2\right] = \\ &= \tfrac{1}{2}\left[\dot{\phi}^2 + \left(\nabla\phi\right)^2 + m^2\phi^2\right]. \end{aligned} \quad \text{(W.63)}$$

3. Starting from $\mathcal{H}$, deduce the Klein–Gordon equation.
Solution: We have

$$\mathcal{H} = \frac{1}{2}\pi^2 - \frac{1}{2}\left(\partial_k\phi\right)\left(\partial^k\phi\right) + \frac{1}{2}m^2\phi^2 = \frac{1}{2}\pi^2 + \frac{1}{2}\left(\partial_k\phi\right)\left(\partial_k\phi\right) + \frac{1}{2}m^2\phi^2. \quad \text{(W.64)}$$

Note, that $\mathcal{H}$ does not depend on $\partial_0\phi$. With (cf. the proceeding section)[81]

[81] Remember

$$\frac{\delta}{\delta\varphi} = \frac{\partial}{\partial\varphi} - \partial_{\mu}\frac{\partial}{\partial\left(\partial_{\mu}\varphi\right)}\ ;\ \frac{\delta}{\delta\pi} = \frac{\partial}{\partial\pi} - \partial_{\mu}\frac{\partial}{\partial\left(\partial_{\mu}\pi\right)} \quad \text{(W.65)}$$

$$\dot{\phi} = \frac{\partial \mathcal{H}}{\partial \pi} - \partial_\mu \frac{\partial \mathcal{H}}{\partial \left(\partial_\mu \pi\right)} \; ; \; \dot{\pi} = -\frac{\partial \mathcal{H}}{\partial \phi} + \partial_\mu \frac{\partial \mathcal{H}}{\partial \left(\partial_\mu \phi\right)} \tag{W.66}$$

follows

$$\begin{aligned} \dot{\phi} = \pi \; ; \; \dot{\pi} &= -m^2\phi + \partial_k \tfrac{\partial \mathcal{H}}{\partial(\partial_k \phi)} = -m^2\phi + \tfrac{1}{2}\partial_k \tfrac{\partial}{\partial(\partial_k \phi)} \left(\partial_l \phi\right)\left(\partial_l \phi\right) = \\ &= -m^2\phi + \partial_k \left(\partial_l \phi\right) \delta_{kl} = -m^2\phi + \partial_k \left(\partial_k \phi\right). \end{aligned} \tag{W.67}$$

This leads to

$$\ddot{\phi} = \dot{\pi} = -m^2\phi + \partial_k \left(\partial_k \phi\right) = -m^2\phi + \nabla^2\phi \tag{W.68}$$

or

$$\partial_\mu \partial^\mu \phi + m^2 \phi = 0. \tag{W.69}$$

4. Prove (W.53).
   Solution: We consider the continuous solution (W.51). With $px = E_{\mathbf{p}}t - \mathbf{px}$ follows

$$\begin{aligned} \phi\left(0, \mathbf{x}\right) &= \int \tfrac{\mathrm{d}^3 p}{\sqrt{(2\pi)^3 2E_{\mathbf{p}}}} \left[a\left(\mathbf{p}\right) e^{i\mathbf{px}} + a^\dagger\left(\mathbf{p}\right) e^{-i\mathbf{px}}\right] \\ \dot{\phi}\left(0, \mathbf{x}\right) &= \int \tfrac{\mathrm{d}^3 p}{\sqrt{(2\pi)^3 2E_{\mathbf{p}}}} \left[-iE_{\mathbf{p}} a\left(\mathbf{p}\right) e^{i\mathbf{px}} + iE_{\mathbf{p}} a^\dagger\left(\mathbf{p}\right) e^{-i\mathbf{px}}\right]. \end{aligned} \tag{W.70}$$

This gives

$$E_{\mathbf{P}}\phi\left(0, \mathbf{x}\right) \pm i\dot{\phi}\left(0, \mathbf{x}\right) = \int \frac{\mathrm{d}^3 p}{\sqrt{(2\pi)^3\, 2E_{\mathbf{p}}}} \left\{ \begin{array}{c} E_{\mathbf{P}}\left[a\left(\mathbf{p}\right) e^{i\mathbf{px}} + a^\dagger\left(\mathbf{p}\right) e^{-i\mathbf{px}}\right] \pm \\ \pm E_{\mathbf{p}}\left[a\left(\mathbf{p}\right) e^{i\mathbf{px}} - a^\dagger\left(\mathbf{p}\right) e^{-i\mathbf{px}}\right] \end{array} \right\}. \tag{W.71}$$

The Fourier transformation of this expression gives with $\int \mathrm{d}^3x \; e^{i\mathbf{Px}} e^{-i\mathbf{px}} = (2\pi)^3\,\delta\left(\mathbf{p} - \mathbf{P}\right)$ and $E_{-\mathbf{P}} = \sqrt{(-\mathbf{P})^2 + m^2} = \sqrt{\mathbf{P}^2 + m^2} = E_{\mathbf{P}}$

$$\int \mathrm{d}^3x\; e^{i\mathbf{Px}} \left[E_{\mathbf{P}}\phi\left(0, \mathbf{x}\right) \pm i\dot{\phi}\left(0, \mathbf{x}\right)\right] = \sqrt{\frac{(2\pi)^3}{2E_{\mathbf{P}}}} \left\{\left(E_{\mathbf{P}} \pm E_{\mathbf{P}}\right) a\left(-\mathbf{P}\right) + \left(E_{\mathbf{P}} \mp E_{\mathbf{P}}\right) a^\dagger\left(\mathbf{P}\right)\right\}. \tag{W.72}$$

Solving for $a_{-\mathbf{P}}$ and $a^\dagger_{\mathbf{P}}$ yields

$$\begin{aligned} \sqrt{(2\pi)^3\, 2E_{\mathbf{P}}}\, a\left(-\mathbf{P}\right) &= \int \mathrm{d}^3x\; e^{i\mathbf{Px}} \left[E_{\mathbf{P}}\phi\left(0, \mathbf{x}\right) + i\dot{\phi}\left(0, \mathbf{x}\right)\right] \\ \sqrt{(2\pi)^3\, 2E_{\mathbf{P}}}\, a^\dagger\left(\mathbf{P}\right) &= \int \mathrm{d}^3x\; e^{i\mathbf{Px}} \left[E_{\mathbf{P}}\phi\left(0, \mathbf{x}\right) - i\dot{\phi}\left(0, \mathbf{x}\right)\right] \end{aligned} \tag{W.73}$$

and finally

$$\begin{aligned} a(\mathbf{P}) &= \frac{1}{\sqrt{(2\pi)^3 2E_\mathbf{P}}} \int \mathrm{d}^3x\ e^{-i\mathbf{Px}} \left[E_\mathbf{P}\phi(0,\mathbf{x}) + i\dot{\phi}(0,\mathbf{x})\right] \\ a^\dagger(\mathbf{P}) &= \frac{1}{\sqrt{(2\pi)^3 2E_\mathbf{P}}} \int \mathrm{d}^3x\ e^{i\mathbf{Px}} \left[E_\mathbf{P}\phi(0,\mathbf{x}) - i\dot{\phi}(0,\mathbf{x})\right]. \end{aligned} \tag{W.74}$$

5. Prove (W.54).
Solution: We consider the continuous case and start from (W.53)

$$\begin{aligned} a(\mathbf{p}) &= \frac{1}{\sqrt{(2\pi)^3 2E_\mathbf{p}}} \int \mathrm{d}^3x\ e^{-i\mathbf{px}} \left(E_\mathbf{p}\phi(0,\mathbf{x}) + i\dot{\phi}(0,\mathbf{x})\right) \\ a^\dagger(\mathbf{p}) &= \frac{1}{\sqrt{(2\pi)^3 2E_\mathbf{p}}} \int \mathrm{d}^3x\ e^{i\mathbf{px}} \left(E_\mathbf{p}\phi(0,\mathbf{x}) - i\dot{\phi}(0,\mathbf{x})\right). \end{aligned} \tag{W.75}$$

It follows

$$a(\mathbf{p})\, a^\dagger(\mathbf{p}') - a^\dagger(\mathbf{p}')\, a(\mathbf{p}) = \frac{1}{\sqrt{(2\pi)^3\, 2E_\mathbf{p}}} \frac{1}{\sqrt{(2\pi)^3\, 2E_{\mathbf{p}'}}} \int \mathrm{d}^3x \int \mathrm{d}^3x'\ e^{-i\mathbf{px}} e^{i\mathbf{p}'\mathbf{x}'}\, \Xi \tag{W.76}$$

with

$$\begin{aligned} \Xi &= \left[\begin{array}{l} \left(E_\mathbf{p}\phi(0,\mathbf{x}) + i\dot{\phi}(0,\mathbf{x})\right)\left(E_{\mathbf{p}'}\phi(0,\mathbf{x}') - i\dot{\phi}(0,\mathbf{x}')\right) - \\ - \left(E_{\mathbf{p}'}\phi(0,\mathbf{x}') - i\dot{\phi}(0,\mathbf{x}')\right)\left(E_\mathbf{p}\phi(0,\mathbf{x}) + i\dot{\phi}(0,\mathbf{x})\right) \end{array}\right] = \\ &= \begin{array}{l} i\dot{\phi}(0,\mathbf{x})\, E_{\mathbf{p}'}\phi(0,\mathbf{x}') - E_\mathbf{p}\phi(0,\mathbf{x})\, i\dot{\phi}(0,\mathbf{x}') - \\ -E_{\mathbf{p}'}\phi(0,\mathbf{x}')\, i\dot{\phi}(0,\mathbf{x}) + i\dot{\phi}(0,\mathbf{x}')\, E_\mathbf{p}\phi(0,\mathbf{x}) \end{array} = \\ &= iE_{\mathbf{p}'}\left[\dot{\phi}(0,\mathbf{x}), \phi(0,\mathbf{x}')\right] + iE_\mathbf{p}\left[\dot{\phi}(0,\mathbf{x}'), \phi(0,\mathbf{x})\right]. \end{aligned} \tag{W.77}$$

With the commutation relations for the field operators (W.49) follows

$$\Xi = iE_{\mathbf{p}'}\left[-i\delta\left(\mathbf{x}-\mathbf{x}'\right)\right] + iE_\mathbf{p}\left[-i\delta\left(\mathbf{x}-\mathbf{x}'\right)\right] = \left(E_{\mathbf{p}'} + E_\mathbf{p}\right)\delta\left(\mathbf{x}-\mathbf{x}'\right). \tag{W.78}$$

With this result, (W.76) reads

$$\begin{aligned} a(\mathbf{p})\, a^\dagger(\mathbf{p}') - a^\dagger(\mathbf{p}')\, a(\mathbf{p}) &= \frac{1}{\sqrt{(2\pi)^3 2E_\mathbf{p}}} \frac{1}{\sqrt{(2\pi)^3 2E_{\mathbf{p}'}}} \int \mathrm{d}^3x \int \mathrm{d}^3x'\ e^{-i\mathbf{px}} e^{i\mathbf{p}'\mathbf{x}'} \left(E_{\mathbf{p}'} + E_\mathbf{p}\right)\delta\left(\mathbf{x}-\mathbf{x}'\right) = \\ &= \frac{1}{\sqrt{(2\pi)^3 2E_\mathbf{p}}} \frac{1}{\sqrt{(2\pi)^3 2E_{\mathbf{p}'}}} \int \mathrm{d}^3x\ e^{-i\mathbf{px}} e^{i\mathbf{p}'\mathbf{x}} \left(E_{\mathbf{p}'} + E_\mathbf{p}\right) = \frac{(2\pi)^3}{\sqrt{(2\pi)^3 2E_\mathbf{p}}} \frac{\delta^{(3)}(\mathbf{p}'-\mathbf{p})}{\sqrt{(2\pi)^3 2E_{\mathbf{p}'}}} \left(E_{\mathbf{p}'} + E_\mathbf{p}\right) \end{aligned} \tag{W.79}$$

or

$$\left[a(\mathbf{p}), a^\dagger(\mathbf{p}')\right] = \delta^{(3)}\left(\mathbf{p}' - \mathbf{p}\right). \tag{W.80}$$

6. Prove (W.58) and/or (W.59).
Solution: We have

$$H = \int \mathrm{d}^3x\ \frac{1}{2}\left[\dot{\phi}^2(x) + (\nabla\phi(x))^2 + m^2\phi^2(x)\right] \tag{W.81}$$

and the continuous form (W.51) of the field operator

$$\phi(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3p}{\sqrt{2E_{\mathbf{p}}}} \left(a(\mathbf{p})\, e^{-ipx} + a^\dagger(\mathbf{p})\, e^{ipx} \right). \tag{W.82}$$

Consider the first summand. With $\partial_t e^{-ipx} = -ip^0 e^{-ipx} = -iE_{\mathbf{p}} e^{-ipx}$ follows

$$\begin{aligned} \int d^3x\, \dot{\phi}^2(x) &= \frac{1}{(2\pi)^3} \int d^3x \int \frac{d^3p}{\sqrt{2E_{\mathbf{p}}}} \int \frac{d^3p'}{\sqrt{2E_{\mathbf{p}'}}} \left[\begin{array}{c} (-ip_0)\left(a(\mathbf{p})\, e^{-ipx} - a^\dagger(\mathbf{p})\, e^{ipx}\right) \cdot \\ \cdot \left(-ip_0'\right)\left(a(\mathbf{p}')\, e^{-ip'x} - a^\dagger(\mathbf{p}')\, e^{ip'x}\right) \end{array} \right] = \\ &= \frac{1}{(2\pi)^3} \int \frac{d^3p}{\sqrt{2E_{\mathbf{p}}}} \int \frac{d^3p'}{\sqrt{2E_{\mathbf{p}'}}} \left(-p_0 p_0'\right) \int d^3x \left[\begin{array}{c} a(\mathbf{p})\, a(\mathbf{p}')\, e^{-ipx} e^{-ip'x} - a^\dagger(\mathbf{p})\, a(\mathbf{p}')\, e^{ipx} e^{-ip'x} - \\ -a(\mathbf{p})\, a^\dagger(\mathbf{p}')\, e^{-ipx} e^{ip'x} + a^\dagger(\mathbf{p})\, a^\dagger(\mathbf{p}')\, e^{ipx} e^{ip'x} \end{array} \right] = \\ &= \int \frac{d^3p}{\sqrt{2E_{\mathbf{p}}}} \int \frac{d^3p'}{\sqrt{2E_{\mathbf{p}'}}} \left[\begin{array}{c} \left(-p_0 p_0'\right) \left\{ a(\mathbf{p})\, a(\mathbf{p}')\, e^{-ip_0x} e^{-ip_0'x} + a^\dagger(\mathbf{p})\, a^\dagger(\mathbf{p}')\, e^{ip_0x} e^{ip_0'x} \right\} \delta(\mathbf{p} + \mathbf{p}') - \\ - \left(-p_0 p_0'\right) \left\{ a^\dagger(\mathbf{p})\, a(\mathbf{p}')\, e^{ip_0x} e^{-ip_0'x} + a(\mathbf{p})\, a^\dagger(\mathbf{p}')\, e^{-ip_0x} e^{ip_0'x} \right\} \delta(\mathbf{p} - \mathbf{p}') \end{array} \right] = \\ &= \int \frac{d^3p}{2E_{\mathbf{p}}} E_{\mathbf{p}}^2 \left[-a(\mathbf{p})\, a(-\mathbf{p})\, e^{-2ip_0x} - a^\dagger(\mathbf{p})\, a^\dagger(-\mathbf{p})\, e^{2ip_0x} + a^\dagger(\mathbf{p})\, a(\mathbf{p}) + a(\mathbf{p})\, a^\dagger(\mathbf{p}) \right]. \end{aligned} \tag{W.83}$$

Note that the energy does not depend on the sign of $\mathbf{p}$ since $E_{\mathbf{p}}^2 = \mathbf{p}^2 + m^2$ and $E_{-\mathbf{p}}^2 = (-\mathbf{p})^2 + m^2 = E_{\mathbf{p}}^2$. Analogously, we arrive at

$$\begin{aligned} \int d^3x\ (\nabla\phi(x))^2 = \int \frac{d^3p}{2E_{\mathbf{p}}} \mathbf{p}^2 \big[& a(\mathbf{p})\, a(-\mathbf{p})\, e^{-2ip_0x} a^\dagger(\mathbf{p})\, a^\dagger(-\mathbf{p})\, e^{2ip_0x} \\ & + a^\dagger(\mathbf{p})\, a(\mathbf{p}) + a(\mathbf{p})\, a^\dagger(\mathbf{p}) \big] \end{aligned} \tag{W.84}$$

and

$$\begin{aligned} m^2 \int d^3x\ \phi^2(x) = \int \frac{d^3p}{2E_{\mathbf{p}}} m^2 \big[& a(\mathbf{p})\, a(-\mathbf{p})\, e^{-2ip_0x} + a^\dagger(\mathbf{p})\, a^\dagger(-\mathbf{p})\, e^{2ip_0x} \\ & + a^\dagger(\mathbf{p})\, a(\mathbf{p}) + a(\mathbf{p})\, a^\dagger(\mathbf{p}) \big]. \end{aligned} \tag{W.85}$$

Adding these three terms brings

$$\begin{aligned} H &= \int d^3x\, \frac{1}{2} \left[\dot{\phi}^2(x) + (\nabla\phi(x))^2 + m^2\phi^2(x) \right] = \\ &= \frac{1}{2} \int \frac{d^3p}{2E_{\mathbf{p}}} \left[\begin{array}{c} \left\{ -E_{\mathbf{p}}^2 + \mathbf{p}^2 + m^2 \right\} \left\{ a(\mathbf{p})\, a(-\mathbf{p})\, e^{-2ip_0x} a^\dagger(\mathbf{p})\, a^\dagger(-\mathbf{p})\, e^{2ip_0x} \right\} + \\ + \left\{ E_{\mathbf{p}}^2 + \mathbf{p}^2 + m^2 \right\} \left\{ a^\dagger(\mathbf{p})\, a(\mathbf{p}) + a(\mathbf{p})\, a^\dagger(\mathbf{p}) \right\} \end{array} \right] = \\ &= \frac{1}{2} \int \frac{d^3p}{2E_{\mathbf{p}}} \left[\left\{ E_{\mathbf{p}}^2 + \mathbf{p}^2 + m^2 \right\} \left\{ a^\dagger(\mathbf{p})\, a(\mathbf{p}) + a(\mathbf{p})\, a^\dagger(\mathbf{p}) \right\} \right] = \\ &= \frac{1}{2} \int d^3p\, E_{\mathbf{p}} \left[a^\dagger(\mathbf{p})\, a(\mathbf{p}) + a(\mathbf{p})\, a^\dagger(\mathbf{p}) \right]. \end{aligned} \tag{W.86}$$

W.5 Quantization of Free Fields, Dirac

W.5.1 No Classical Spinor Field

In the last section, we have applied the method of canonical quantization to the Klein–Gordon field. However, for the Dirac field, this approach does not work. This is due to the fact that the Dirac equation describes electrons and positrons[82] which have spin $\frac{1}{2}$. In other words, these particles are fermions. As such, they obey the Pauli exclusion principle stating that two fermions cannot occupy the same state.

In contrast, bosons as for instance photons can populate the same state without restriction. Thus, they reinforce one another and can produce a macroscopic field; in case of photons the electromagnetic field. This mechanism is not accessible for fermions. This means there is no macroscopic spinor field and, therefore, no corresponding $\mathcal{L}$ or $\mathcal{H}$, let alone Poisson brackets which we could second-quantize via the canonical quantization.

So we have to choose another approach. First, we assume that the Dirac equation is also the underlying equation for the quantized field, as we did in the case of the Klein–Gordon field. Next, we search for an Lagrangian $\mathcal{L}$ which reproduces the Dirac equation. Equipped with this information, we can calculate the conjugated momentum, the Hamiltonian density $\mathcal{H}$ and the Hamilton function H. Then, parallelizing the Klein–Gordon case, we insert into H the free solutions of the Dirac equation which are composed of free waves e^{-ikx} and e^{ikx} with amplitudes b and d. Again, we change these quantities into operators. However, we have no Poisson brackets and have to look for suitable commutation relations for b and d which are compatible with theoretical considerations and, first of all, with experimental results.

W.5.2 Lagrangian, Conjugated Momentum, Hamiltonian

As it turns out, an appropriate Lagrangian is[83]

$$\mathcal{L} = \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi \ ; \ \psi = \psi(x) . \tag{W.87}$$

The 4-spinor ψ is complex and has real and imaginary parts. This means that we can regard the two fields ψ and $\bar{\psi}$ as independent.

The conjugated fields for ψ and $\bar{\psi}$ are given by[84]

---

[82]Of course, the DE is valid for all particles with spin $1/2$, e.g. also for muons or tauons and their antiparticles.

[83]Remember $\bar{\psi} = \psi^{\dagger}\gamma^{0}$ where $\psi^{\dagger}$ is the hermitian adjoint and $\bar{\psi}$ the (Dirac) adjoint.

[84]For the variational derivative $\frac{\delta}{\delta f}$ see Vol. 1, App. T.3.

$$\text{for } \psi\text{: } \pi_\psi = \frac{\delta\mathcal{L}}{\delta\left(\partial_0\psi\right)} = i\bar{\psi}\gamma^0 \text{ ; for } \bar{\psi}\text{: } \pi_{\bar{\psi}} = \frac{\delta\mathcal{L}}{\delta\left(\partial_0\bar{\psi}\right)} = 0. \tag{W.88}$$

The Hamiltonian density follows from $\mathcal{H} = \pi_\psi\partial_0\psi - \mathcal{L}$ as

$$\mathcal{H} = \pi_\psi\partial_0\psi - \mathcal{L} = i\bar{\psi}\gamma^0\partial_0\psi - \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi = -\bar{\psi}\left(i\gamma^k\partial_k - m\right)\psi. \tag{W.89}$$

This expression may be simplified taking into account the Dirac equation in the form $i\gamma^0\partial_0\psi = -\left(i\gamma^k\partial_k - m\right)\psi$, yielding

$$\mathcal{H} = \bar{\psi}i\gamma^0\partial_0\psi = i\psi^\dagger\partial_0\psi. \tag{W.90}$$

Finally, the Hamilton function H reads

$$H = i\int \mathrm{d}^3x\ \psi^\dagger\partial_0\psi. \tag{W.91}$$

W.5.3 The Free Solutions

We now invoke the free solutions of the Dirac equation (cf. Appendix U, Vol. 1). The continuous version reads

$$\psi(x) = \frac{1}{(2\pi)^{3/2}}\int \mathrm{d}^3p\sqrt{\frac{m}{E_\mathbf{p}}}\sum_r\left(b_r(\mathbf{p})\,u_r(\mathbf{p})\,e^{-ipx} + \mathrm{d}_r^\dagger(\mathbf{p})\,w_r(\mathbf{p})\,e^{ipx}\right) \tag{W.92}$$

and the discrete version is obtained by the change $\frac{1}{(2\pi)^{3/2}}\int \mathrm{d}^3p \rightarrow \frac{1}{\sqrt{V}}\sum_\mathbf{p}$ (cf. Appendix T, Vol. 1).[85] $r = 1, 2$ denotes the spin directions.

W.5.3.1 Properties of the Spinors u_r and w_r

In further considerations, we make use of the properties of the spinors u_r and w_r which we will discuss now.

Note that there is a (minor) difference between (W.92) and the free solutions as formulated in Appendix U, Vol. 1. There we have written v_r instead of w_r, namely

$$\text{in Vol. 1: } \psi(x) = \frac{1}{(2\pi)^{3/2}}\int \mathrm{d}^3p\sqrt{\frac{m}{E_\mathbf{p}}}\sum_r\left(b_r(\mathbf{p})\,u_r(\mathbf{p})\,e^{-ipx} + \mathrm{d}_r^\dagger(\mathbf{p})\,v_r(\mathbf{p})\,e^{ipx}\right) \tag{W.93}$$

with

[85] As in the Klein–Gordon case, we use for convenience the same symbols $b(\mathbf{p})$ and $d(\mathbf{p})$ in the discrete and the continuous case.

$$u_r(\mathbf{p}) = \begin{pmatrix} \left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2} \chi_r \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(m+E_\mathbf{p}))^{1/2}} \chi_r \end{pmatrix} \; ; \; v_r(\mathbf{p}) = \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(m+E_\mathbf{p}))^{1/2}} \chi_r \\ \left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2} \chi_r \end{pmatrix} \tag{W.94}$$

where the χ_r are the 2-spinors

$$\chi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \; ; \; \chi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{W.95}$$

However, with regard to further considerations, it is advantageous (and therefore quite common) to replace v by w defined by

$$w_r(\mathbf{p}) = \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(E_\mathbf{p}+m))^{1/2}} i\sigma_2 \chi_r \\ \left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2} i\sigma_2 \chi_r \end{pmatrix} \text{ or } w_1(\mathbf{p}) = v_2(\mathbf{p}) \; ; \; w_2(\mathbf{p}) = -v_1(\mathbf{p}). \tag{W.96}$$

In doing so, we will be able below to cast certain relations in a simpler manner.

For the pair u, v we have derived in Appendix U, Vol. 1 the following orthogonality relations:

$$\begin{aligned} \bar{u}_r(\mathbf{p}) u_{r'}(\mathbf{p}) &= \delta_{rr'} \; ; \; \bar{u}_r(\mathbf{p}) v_{r'}(\mathbf{p}) = 0 \\ \bar{v}_r(\mathbf{p}) v_{r'}(\mathbf{p}) &= -\delta_{rr'} \; ; \; \bar{v}_r(\mathbf{p}) u_{r'}(\mathbf{p}) = 0. \end{aligned} \tag{W.97}$$

Using w instead of v, these relations translate into

$$\begin{aligned} \bar{u}_r(\mathbf{p}) u_{r'}(\mathbf{p}) &= \delta_{rr'} \; ; \; \bar{u}_r(\mathbf{p}) w_{r'}(\mathbf{p}) = 0 \\ \bar{w}_r(\mathbf{p}) w_{r'}(\mathbf{p}) &= -\delta_{rr'} \; ; \; \bar{w}_r(\mathbf{p}) u_{r'}(\mathbf{p}) = 0. \end{aligned} \tag{W.98}$$

In addition we have

$$\begin{aligned} \bar{u}_r(\mathbf{p}) \gamma^0 u_{r'}(\mathbf{p}) &= \tfrac{E_\mathbf{p}}{m} \delta_{rr'} \; ; \; \bar{u}_r(-\mathbf{p}) \gamma^0 w_{r'}(\mathbf{p}) = 0 \\ \bar{w}_r(\mathbf{p}) \gamma^0 w_{r'}(\mathbf{p}) &= \tfrac{E_\mathbf{p}}{m} \delta_{rr'} \; ; \; \bar{w}_r(-\mathbf{p}) \gamma^0 u_{r'}(\mathbf{p}) = 0. \end{aligned} \tag{W.99}$$

As in the Klein–Gordon case, we adopt the free solutions (W.92) as field operators, i.e., we interpret now the 'amplitudes' $b_r(\mathbf{p})$ and $\mathrm{d}_r(\mathbf{p})$ as operators. Apparently, there is a new feature in comparison with the Klein–Gordon case: we have now *two* types of operators.[86] In accordance with the considerations in Appendix U, Vol. 1 concerning the electron and its antiparticle, we regard them as acting on *electrons* and *positrons*: $b_r(\mathbf{p})$ and $b_r^\dagger(\mathbf{p})$ destroy and create an electron, $\mathrm{d}_r(\mathbf{p})$ and $\mathrm{d}_r^\dagger(\mathbf{p})$ destroy and create a positron. Note that at the current state of our considerations, this is a guess or assumption which has to prove to be true.

[86]This indicates that the particles described differ from their antiparticles.

W.5.4 Energy

Since in the Dirac case we have no Poisson brackets from a macroscopic system, we have to proceed in another way to get information about the commutation rules for the operators $b_r(\mathbf{p})$ and $\mathrm{d}_r(\mathbf{p})$. To this end, we first determine the Hamilton function (or energy) (W.91) in terms of these operators. On this basis, we can discuss appropriate commutation rules, i.e., an appropriate way to quantize the Dirac field.

Inserting the free solutions (W.92) into (W.91), i.e., $H = i\int \mathrm{d}^3x\ \psi^\dagger\partial_0\psi$, leads to the following expressions for the continuous version (see exercises):

$$H = i\int \mathrm{d}^3x\ \psi^\dagger\partial_0\psi = \int \mathrm{d}^3p\ E_\mathbf{p}\sum_r \left[b_r^\dagger(\mathbf{p})\, b_r(\mathbf{p}) - \mathrm{d}_r(\mathbf{p})\,\mathrm{d}_r^\dagger(\mathbf{p})\right] \tag{W.100}$$

whereas the discrete version is obtained by the replacement $\int \mathrm{d}^3p \to \sum_\mathbf{p}$.

W.5.5 Interpretation of $b_r(\mathbf{p})$ and $d_r(\mathbf{p})$, Commutation Relations, Pauli Principle, Number Operator

On the basis of the expressions (W.100) , we will now try to find out which commutation rules are appropriate for the Dirac system. Remember our assumption as presented above that $b_r(\mathbf{p})$ and $b_r^\dagger(\mathbf{p})$ destroy and create an electron, $\mathrm{d}_r(\mathbf{p})$ and $\mathrm{d}_r^\dagger(\mathbf{p})$ destroy and create a positron. This means, for example, that applying the operator $b_r^\dagger(\mathbf{p})$ to the vacuum state $|0\rangle$ creates an electron with quantum numbers r and $\mathbf{p}$; analogously with $\mathrm{d}_r^\dagger(\mathbf{p})$ for positrons.

W.5.5.1 Anticommutation Relations

One could postulate that the operators $b_r(\mathbf{p})$ and $\mathrm{d}_r(\mathbf{p})$ obey similar commutation relations as in the Klein–Gordon case, i.e., $\left[b_r(\mathbf{p}), b_{r'}^\dagger(\mathbf{p}')\right] \sim \delta_{rr'}\delta_{\mathbf{pp}'}$. But doing so leads to an unviable theory with a lot of inconsistencies[87] which, above all, would not reproduce the fact that electrons and positrons are fermions.

As a matter of fact, one has to introduce *anticommutation relations*.[88] They read

$$\left\{b_r(\mathbf{p}), b_{r'}^\dagger(\mathbf{p}')\right\} = \delta_{rr'}\delta\left(\mathbf{p},\mathbf{p}'\right)\ \ ;\ \ \left\{\mathrm{d}_r(\mathbf{p}), d_{r'}^\dagger(\mathbf{p}')\right\} = \delta_{rr'}\delta\left(\mathbf{p},\mathbf{p}'\right). \tag{W.101}$$

All other anticommutators between two of these operators vanish,

[87]For instance, with commutation rules as with bosons, the energy would not be bounded from below.

[88]Note that this step is not mandatory or logically without alternative at this state of affairs and has to prove itself. Just a reminder: $\{a, b\} = ab + ba$.

$$\left\{b_r(\mathbf{p}), b_{r'}\left(\mathbf{p}'\right)\right\} = \left\{\mathrm{d}_r(\mathbf{p}), b_{r'}\left(\mathbf{p}'\right)\right\} = \left\{b_r(\mathbf{p}), \mathrm{d}^\dagger_{r'}\left(\mathbf{p}'\right)\right\} = \cdots = 0. \quad \text{(W.102)}$$

W.5.5.2 Pauli Principle

Note that the anticommutation relations (W.101) and (W.102) guarantee that we describe fermions. For instance, from (W.102) follows $\{b_r(\mathbf{p}), b_r(\mathbf{p})\} = 2b_r(\mathbf{p})\, b_r(\mathbf{p}) = 0$ and analogously for $b_r^\dagger(\mathbf{p})$, $\mathrm{d}_r(\mathbf{p})$ and $\mathrm{d}_r^\dagger(\mathbf{p})$, i.e.,

$$b_r(\mathbf{p})\, b_r(\mathbf{p}) = 0\ ;\ b_r^\dagger(\mathbf{p})\, b_r^\dagger(\mathbf{p}) = 0\ ;\ \mathrm{d}_r(\mathbf{p})\, \mathrm{d}_r(\mathbf{p}) = 0\ ;\ \mathrm{d}_r^\dagger(\mathbf{p})\, \mathrm{d}_r^\dagger(\mathbf{p}) = 0. \quad \text{(W.103)}$$

Assume that we create e.g. an electron with quantum numbers $r\mathbf{p}$, i.e., we have $b_r^\dagger(\mathbf{p})\,|0\rangle$. Applying once more $b_r^\dagger(\mathbf{p})$ would, in the Klein–Gordon case, produce a second particle with the same quantum numbers. But in the Dirac case, this has to be forbidden due to the Pauli exclusion principle. Indeed, we have from (W.103)

$$b_r^\dagger(\mathbf{p})\, b_r^\dagger(\mathbf{p})\,|0\rangle = 0. \quad \text{(W.104)}$$

Thus, each state is either empty or simply occupied. Obviously, a commutation relation like $\left[b_r^\dagger(\mathbf{p}), b_{r'}^\dagger\left(\mathbf{p}'\right)\right] \sim \delta_{rr'}\delta_{\mathbf{pp}'}$ would not produce this behavior.

W.5.5.3 Number Operator

Parallelizing the Klein–Gordon case, we define a number operator for electrons and positrons by

$$N_{e,r\mathbf{p}} = b_r^\dagger(\mathbf{p})\, b_r(\mathbf{p})\ \ ;\ N_{p,r\mathbf{p}} = \mathrm{d}_r^\dagger(\mathbf{p})\, \mathrm{d}_r(\mathbf{p}) \quad \text{(W.105)}$$

which give us the number of electrons and positrons with quantum numbers r and $\mathbf{p}$.

Note that electrons and positrons are now on the same level, and we do not have any more the situation of one single electron against an infinite sea of holes, i.e., positrons.

We want to reproduce the Pauli principle once again, this times by using the number operator. We apply $N_{e,r\mathbf{p}}$ onto a state $|\psi\rangle$ which consists of n electrons with quantum numbers r and $\mathbf{p}$, i.e., $N_{e,r\mathbf{p}}\,|\psi\rangle = n\,|\psi\rangle$. Due to the anticommutation relations, we have

$$\begin{aligned} N_{e,r\mathbf{p}}N_{e,r\mathbf{p}} &= b_r^\dagger(\mathbf{p})\, b_r(\mathbf{p})\, b_r^\dagger(\mathbf{p})\, b_r(\mathbf{p}) = b_r^\dagger(\mathbf{p})\left[1 - b_r^\dagger(\mathbf{p})\, b_r(\mathbf{p})\right] b_r(\mathbf{p}) = \\ &= b_r^\dagger(\mathbf{p})\, b_r(\mathbf{p}) - b_r^\dagger(\mathbf{p})\, b_r^\dagger(\mathbf{p})\, b_r(\mathbf{p})\, b_r(\mathbf{p}) = b_r^\dagger(\mathbf{p})\, b_r(\mathbf{p}) = N_{e,r\mathbf{p}} \end{aligned} \quad \text{(W.106)}$$

and it follows

$$N_{e,r\mathbf{p}}N_{e,r\mathbf{p}}\,|\psi\rangle = N_{e,r\mathbf{p}}\,|\psi\rangle \rightarrow n^2\,|\psi\rangle = n\,|\psi\rangle \rightarrow n = 0 \text{ or } n = 1. \quad \text{(W.107)}$$

Thus, $N_{e,r\mathbf{p}}$ has the eigenvalues 0 and 1. In other words, we have again found the Pauli principle that for fermions each state is either empty or simply occupied.

W.5.6 Again Infinities

Since we want to write H as a combination of number operators, we invoke the anticommutation rules and cast

$$H = \sum_{\mathbf{p},r} E_{\mathbf{p}} \left[b_r^\dagger(\mathbf{p})\, b_r(\mathbf{p}) - \mathrm{d}_r(\mathbf{p})\, \mathrm{d}_r^\dagger(\mathbf{p}) \right] \tag{W.108}$$

into the form

$$H = \sum_{\mathbf{p},r} E_{\mathbf{p}} \left[b_r^\dagger(\mathbf{p})\, b_r(\mathbf{p}) + \mathrm{d}_r^\dagger(\mathbf{p})\, \mathrm{d}_r(\mathbf{p}) - 1 \right]. \tag{W.109}$$

As in the Klein–Gordon case, we here have the problem of an infinite contribution $-\sum_{\mathbf{p},r} E_{\mathbf{p}}$ even if there are no particles.

It would be nice if we could simply neglect the infinite zero point energy to get around this conceptual difficulty and could write

$$H = \sum_{\mathbf{p},r} E_{\mathbf{p}} \left[b_r^\dagger(\mathbf{p})\, b_r(\mathbf{p}) + \mathrm{d}_r^\dagger(\mathbf{p})\, \mathrm{d}_r(\mathbf{p}) \right] = \sum_{\mathbf{p},r} E_{\mathbf{p}} \left[N_{e,r\mathbf{p}} + N_{p,r\mathbf{p}} \right]. \tag{W.110}$$

Indeed, this is the final form of the Hamilton operator. A thorough discussion of this delicate point is found below in the section 'Operator ordering'.

W.5.7 Anticommutators for Field Operators

As said above, we cannot use the scheme of canonical quantization for the Dirac system, since there is no macroscopic spinor theory and, correspondingly, there are no Poisson brackets which we could quantize. But now, in possession of the appropriate commutators (W.101) for fermions, we can ask which relation may replace the (quantized) Poisson brackets. It is clear that there have to be differences with regard to e.g. the Klein–Gordon field, if only because we have anticommutators instead of commutators and 4-spinors instead of a scalar field. We will discuss this question only very briefly and without detailed calculations.

We sketch the approach in all brevity. One starts with the free solutions as given in (W.92), i.e.,

$$\psi(x) = \sum_{\mathbf{p},r} \sqrt{\frac{m}{VE_{\mathbf{p}}}} \left(b_r(\mathbf{p})\, u_r(\mathbf{p})\, e^{-ipx} + \mathrm{d}_r^\dagger(\mathbf{p})\, w_r(\mathbf{p})\, e^{ipx} \right) \tag{W.111}$$

for the discrete case (the continuous case runs analogously). By means of Fourier transformation, one can solve this equation for $b_r(\mathbf{p})\, u_r(p)$ and $d_r^\dagger(\mathbf{p})\, w_r(p)$. Invoking the anticommutation relations for $b_r(\mathbf{p})$ and $d_r(\mathbf{p})$ and the orthogonality relations for $u_r(p)$ and $w_r(p)$, one obtains after some calculations the anticommutation relations for the field operators

$$\begin{aligned} \left\{\psi_\alpha(t,\mathbf{x}), \psi_\beta^\dagger\left(t,\mathbf{x}'\right)\right\} &= \delta_{\alpha\beta}\delta^{(3)}(\mathbf{x}-\mathbf{x}') \\ \left\{\psi_\alpha(t,\mathbf{x}), \psi_\beta\left(t,\mathbf{x}'\right)\right\} = \left\{\psi_\alpha^\dagger(t,\mathbf{x}), \psi_\beta^\dagger\left(t,\mathbf{x}'\right)\right\} &= 0 \end{aligned} \tag{W.112}$$

where $\alpha = 1, \ldots, 4$ and $\beta = 1, \ldots, 4$ indicate the components of the spinors. Note that these are equal-time relations as in the Klein–Gordon case. The relations are formulated in terms of the hermitian adjoint $\psi^\dagger$ and not the Dirac adjoint $\bar{\psi}$. With $\bar{\psi}$, we have for instance

$$\left\{\psi_\alpha(t,\mathbf{x}), \bar{\psi}_\beta\left(t,\mathbf{x}'\right)\right\} = \gamma^0_{\alpha\beta}\delta^{(3)}(\mathbf{x}-\mathbf{x}'). \tag{W.113}$$

Of course, one can postulate (W.112) out of the blue and then derive the anticommutation relations (W.101) for the creation/annihilation operators. But it seems hard to see, especially for beginners, why relations (W.112) should have precisely the form they have.

W.5.8 Conclusion

With the relations (W.101), we have quantized the Dirac field; the Hamiltonian is given in (W.110). The operators $b_r^\dagger(\mathbf{p})$ and $b_r(\mathbf{p})$ create and annihilate an electron and the operators $d_r^\dagger(\mathbf{p})$ and $d_r(\mathbf{p})$ create and annihilate a positron, both with quantum numbers r and $\mathbf{p}$. The states are living in a Fock space, for instance $|0\rangle$ or $d_r^\dagger(\mathbf{p})\,|0\rangle$ or $b_{r_1}^\dagger(\mathbf{p}_1)\, d_{r_2}^\dagger(\mathbf{p}_2)\, d_{r_3}^\dagger(\mathbf{p}_3)\,|0\rangle$ (an electron with $r_1\mathbf{p}_1$, a positron with $r_2\mathbf{p}_2$, a positron with $r_3\mathbf{p}_3$). Concerning the spin, we have the following facts: $b_r^\dagger(\mathbf{p}=\mathbf{0})$ / $b_r(\mathbf{p}=\mathbf{0})$ with $r = 1$ ($r = 2$) creates/annihilates a stationary electron with spin $s_z = \frac{1}{2}$ ($s_z = -\frac{1}{2}$); analogously with $d_r^\dagger(\mathbf{p}=\mathbf{0})$ / $d_r(\mathbf{p}=\mathbf{0})$ for positrons. In general, $b_r^\dagger(\mathbf{p})$ or $d_r^\dagger(\mathbf{p})$ creates an electron or positron with momentum $\mathbf{p}$ which in its rest system has the spin $\frac{1}{2}$ and $-\frac{1}{2}$ for $r = 1$ and $r = 2$.

Some normalized particle states with their energy eigenvalues are given in Table W.2.

In sum, we have now an physical meaningful picture; $\psi(x)$ is not a state, but a field operator, creating and annihilating particles, i.e., electrons and positrons. These two types of particles are now on equal footing, and we do not need anymore an infinite sea of positrons in order to describe one electron as in the Dirac theory of Appendix U, Vol. 1. In other words, we have left the one-particle theories, and can describe arbitrary numbers of particles.

Table W.2 Table of simplest states of the Dirac field

State		Energy
Vacuum	$\|0\rangle$	0
One electron	$b_r^{\dagger}(\mathbf{p})\|0\rangle$	$E_{\mathbf{p}}$
One positron	$\mathrm{d}_r^{\dagger}(\mathbf{p})\|0\rangle$	$E_{\mathbf{p}}$
One electron, one positron	$b_r^{\dagger}(\mathbf{p})\,\mathrm{d}_r^{\dagger}(\mathbf{p})\|0\rangle$	$2E_{\mathbf{p}}$
Two different electrons	$b_{r'}^{\dagger}(\mathbf{p}')\,b_r^{\dagger}(\mathbf{p})\|0\rangle$	$E_{\mathbf{p}'}+E_{\mathbf{p}}$; $\mathbf{p}'\neq\mathbf{p}$ and /or $r'\neq r$
Two identical electrons	$b_r^{\dagger}(\mathbf{p})\,b_r^{\dagger}(\mathbf{p})\|0\rangle$	0

In addition, there are no problems with negative energies. The Hamiltonian (W.110) simply does not allow for them.

All fits nicely, and the only weak spot, if one may say, is the infinite contribution to the Hamiltonian in (W.109) and its negligence in (W.110). For the time being, we must accept this as the way nature works.

W.5.9 *Exercises and Solutions*

1. Show that the Euler–Lagrange equations for ψ and $\bar{\psi}$ reproduce the Dirac equation.
   Solution: We have

$$\frac{\partial\mathcal{L}}{\partial\bar{\psi}}-\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\bar{\psi}\right)}\right)=0\ \text{ and }\ \frac{\partial\mathcal{L}}{\partial\psi}-\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi\right)}\right). \tag{W.114}$$

   The Lagrangian being $\mathcal{L}=\bar{\psi}\left(i\gamma^\mu\partial_\mu-m\right)\psi$, we have

$$\frac{\partial\mathcal{L}}{\partial\bar{\psi}}=\left(i\gamma^\mu\partial_\mu-m\right)\psi\ ;\ \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\bar{\psi}\right)}=0\Longrightarrow\left(i\gamma^\mu\partial_\mu-m\right)\psi=0 \tag{W.115}$$

   and

$$\frac{\partial\mathcal{L}}{\partial\psi}=-m\bar{\psi}\ ;\ \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi\right)}=\frac{\partial\bar{\psi}i\gamma^\nu\partial_\nu\psi}{\partial\left(\partial_\mu\psi\right)}=\bar{\psi}i\gamma^\nu\delta_{\nu\mu}\Longrightarrow-m\bar{\psi}-\partial_\mu\bar{\psi}i\gamma^\mu=0. \tag{W.116}$$

   We see that (W.115) yields the Dirac equation and (W.116) its adjoint.
2. Write down $\psi^{\dagger}(x)$ and $\bar{\psi}(x)$.
   Solution: It is

$$\psi^{\dagger}(x) = \sum_{\mathbf{p},r=1,2} \left(\frac{m}{VE_{\mathbf{p}}}\right)^{1/2} \left(b_r^{\dagger}(\mathbf{p})\, u_r^{\dagger}(\mathbf{p})\, e^{ipx} + \mathrm{d}_r(\mathbf{p})\, w_r^{\dagger}(\mathbf{p})\, e^{-ipx}\right). \tag{W.117}$$

With $\bar{\psi} = \psi^{\dagger}\gamma^0$ follows:

$$\begin{aligned}\bar{\psi}(x) = \psi^{\dagger}(x)\,\gamma^0 = \textstyle\sum_{\mathbf{p},r=1,2} \left(\frac{m}{VE_{\mathbf{p}}}\right)^{1/2} \left(b_r^{\dagger}(\mathbf{p})\, u_r^{\dagger}(\mathbf{p})\,\gamma^0 e^{ipx} + \mathrm{d}_r(\mathbf{p})\, w_r^{\dagger}(\mathbf{p})\,\gamma^0 e^{-ipx}\right) = \\ = \textstyle\sum_{\mathbf{p},r=1,2} \left(\frac{m}{VE_{\mathbf{p}}}\right)^{1/2} \left(b_r^{\dagger}(\mathbf{p})\, \bar{u}_r(\mathbf{p})\, e^{ipx} + \mathrm{d}_r(\mathbf{p})\, \bar{w}_r(\mathbf{p})\, e^{-ipx}\right).\end{aligned} \tag{W.118}$$

3. Write down explicitly u_r in (W.94) and w_r in (W.96) for $r = 1, 2$.
   Solution:

$$\begin{aligned}
u_1(\mathbf{p}) &= \left(\frac{E_{\mathbf{p}}+m}{2m}\right)^{1/2} \begin{pmatrix} \chi_1 \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E_{\mathbf{p}}+m}\chi_1 \end{pmatrix} = \left(\frac{E_{\mathbf{p}}+m}{2m}\right)^{1/2} \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E_{\mathbf{p}}+m} \\ \frac{p_x+ip_y}{E_{\mathbf{p}}+m} \end{pmatrix} \\
u_2(\mathbf{p}) &= \left(\frac{E_{\mathbf{p}}+m}{2m}\right)^{1/2} \begin{pmatrix} \chi_2 \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E_{\mathbf{p}}+m}\chi_2 \end{pmatrix} = \left(\frac{E_{\mathbf{p}}+m}{2m}\right)^{1/2} \begin{pmatrix} 0 \\ 1 \\ \frac{p_x-ip_y}{E_{\mathbf{p}}+m} \\ \frac{-p_z}{E_{\mathbf{p}}+m} \end{pmatrix} \\
w_1(\mathbf{p}) &= -\left(\frac{E_{\mathbf{p}}+m}{2m}\right)^{1/2} \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E_{\mathbf{p}}+m} i\sigma_2\chi_1 \\ i\sigma_2\chi_1 \end{pmatrix} = -\left(\frac{E_{\mathbf{p}}+m}{2m}\right)^{1/2} \begin{pmatrix} \frac{-p_x+ip_y}{E_{\mathbf{p}}+m} \\ \frac{p_z}{E_{\mathbf{p}}+m} \\ 0 \\ -1 \end{pmatrix} \\
w_2(\mathbf{p}) &= -\left(\frac{E_{\mathbf{p}}+m}{2m}\right)^{1/2} \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E_{\mathbf{p}}+m} i\sigma_2\chi_2 \\ i\sigma_2\chi_2 \end{pmatrix} = -\left(\frac{E_{\mathbf{p}}+m}{2m}\right)^{1/2} \begin{pmatrix} \frac{p_z}{E_{\mathbf{p}}+m} \\ \frac{p_x+ip_y}{E_{\mathbf{p}}+m} \\ 1 \\ 0 \end{pmatrix}.
\end{aligned} \tag{W.119}$$

4. Write down explicitly $\bar{w}_r$ for $r = 1, 2$.
   Solution: It is $\bar{w}_r = w_r^{\dagger}\gamma_0$ and

$$w_r(\mathbf{p}) = \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{\left(2m\left(m+E_{\mathbf{p}}\right)\right)^{1/2}} i\sigma_2\chi_r \\ \left(\frac{E_{\mathbf{p}}+m}{2m}\right)^{1/2} i\sigma_2\chi_r \end{pmatrix}. \tag{W.120}$$

With $(i\sigma_2)^{\dagger} = -i\sigma_2$ follows

$$\bar{w}_r(\mathbf{p}) = \left(-\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{\left(2m\left(m+E_{\mathbf{p}}\right)\right)^{1/2}} i\sigma_2\chi_r \;\; \left(\frac{E_{\mathbf{p}}+m}{2m}\right)^{1/2} i\sigma_2\chi_r\right). \tag{W.121}$$

5. By means of the explicit expressions for u_r and w_r, check equations (W.98). Solution: We calculate $\bar{w}_r\,(p)\,w_{r'}\,(p)$ and $\bar{w}_r\,(p)\,u_{r'}\,(p)$. We have[89]

$$\begin{aligned}\bar{w}_r(\mathbf{p})w_{r'}(\mathbf{p}) &= \left(-\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(E_\mathbf{p}+m))^{1/2}}\chi_r^\dagger i\sigma_2 \quad \left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2}\chi_r^\dagger i\sigma_2\right)\begin{pmatrix}\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(E_\mathbf{p}+m))^{1/2}}i\sigma_2\chi_{r'}\\ \left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2} i\sigma_2\chi_{r'}\end{pmatrix}=\\ &= -\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(E_\mathbf{p}+m))^{1/2}}\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(E_\mathbf{p}+m))^{1/2}}\left(\chi_r^\dagger i\sigma_2\right)(i\sigma_2\chi_{r'})+\left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2}\left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2}\left(\chi_r^\dagger i\sigma_2\right)(i\sigma_2\chi_{r'})=\\ &=\left[-\frac{\mathbf{p}^2}{2m(E_\mathbf{p}+m)}+\frac{E_\mathbf{p}+m}{2m}\right]\chi_r^\dagger i\sigma_2 i\sigma_2\chi_{r'}=\frac{(E_\mathbf{p}+m)^2-\mathbf{p}^2}{2m(E_\mathbf{p}+m)}\chi_r^\dagger(-1)\chi_{r'}=\chi_r^\dagger(-1)\chi_{r'}=--\delta_{rr'}\end{aligned} \tag{W.122}$$

and

$$\begin{aligned}\bar{w}_r(\mathbf{p})u_{r'}(\mathbf{p}) &= \left(-\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(E_\mathbf{p}+m))^{1/2}}\chi_r^\dagger i\sigma_2 \quad \left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2}\chi_r^\dagger i\sigma_2\right)\begin{pmatrix}\left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2}\chi_r\\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(m+E_\mathbf{p}))^{1/2}}\chi_r\end{pmatrix}=\\ &= -\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(E_\mathbf{p}+m))^{1/2}}\chi_r^\dagger i\sigma_2\left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2}\chi_r+\left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2}\chi_r^\dagger i\sigma_2\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(m+E_\mathbf{p}))^{1/2}}\chi_r=\\ &=\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{(2m(E_\mathbf{p}+m))^{1/2}}\left(\frac{E_\mathbf{p}+m}{2m}\right)^{1/2}\left[-\chi_r^\dagger i\sigma_2\chi_r+\chi_r^\dagger i\sigma_2\chi_r\right]=0.\end{aligned} \tag{W.123}$$

6. By means of the explicit expressions for u_r and w_r, check equations (W.99). Solution: By proxy, we calculate $\bar{u}_1(\mathbf{p})\gamma^0 u_{2'}(\mathbf{p})$, $\bar{u}_1(\mathbf{p})\gamma^0 u_{1'}(\mathbf{p})$ and $\bar{u}_r(-\mathbf{p})\gamma^0 w_{r'}(\mathbf{p})$. For $u_1^\dagger(\mathbf{p})u_2(\mathbf{p})$, we have

$$\begin{aligned}\bar{u}_1(\mathbf{p})\gamma^0 u_{2'}(\mathbf{p}) = u_1^\dagger(\mathbf{p})u_2(\mathbf{p}) &= \frac{E_\mathbf{p}+m}{2m}\begin{pmatrix}1\\0\\ \frac{p_z}{E_\mathbf{p}+m}\\ \frac{p_x-ip_y}{E_\mathbf{p}+m}\end{pmatrix}^T\begin{pmatrix}0\\1\\ \frac{p_x-ip_y}{E_\mathbf{p}+m}\\ \frac{-p_z}{E_\mathbf{p}+m}\end{pmatrix}=\\ &=\frac{E_\mathbf{p}+m}{2m}\left(0+0+\frac{p_z}{E_\mathbf{p}+m}\frac{p_x-ip_y}{E_\mathbf{p}+m}+\frac{p_x-ip_y}{E_\mathbf{p}+m}\frac{-p_z}{E_\mathbf{p}+m}\right)=\\ &=\frac{E_\mathbf{p}+m}{2m}\frac{p_z}{E_\mathbf{p}+m}\left(\frac{p_x-ip_y}{E_\mathbf{p}+m}-\frac{p_x-ip_y}{E_\mathbf{p}+m}\right)=0.\end{aligned} \tag{W.124}$$

$\bar{u}_1(\mathbf{p})\gamma^0 u_{1'}(\mathbf{p}) = u_1^\dagger(\mathbf{p})u_1(\mathbf{p})$ is given by

$$\begin{aligned}u_1^\dagger(\mathbf{p})u_1(\mathbf{p}) &= \frac{E_\mathbf{p}+m}{2m}\begin{pmatrix}1\\0\\ \frac{p_z}{E_\mathbf{p}+m}\\ \frac{p_x+ip_y}{E_\mathbf{p}+m}\end{pmatrix}^T\begin{pmatrix}1\\0\\ \frac{p_z}{E_\mathbf{p}+m}\\ \frac{p_x-ip_y}{E_\mathbf{p}+m}\end{pmatrix}=\frac{E_\mathbf{p}+m}{2m}\left(1+0+\frac{p_z^2}{(E_\mathbf{p}+m)^2}+\frac{p_x^2+p_y^2}{(E_\mathbf{p}+m)^2}\right)=\\ &=\frac{E_\mathbf{p}+m}{2m}\left(1+\frac{E_\mathbf{p}^2-m^2}{(E_\mathbf{p}+m)^2}\right)=\frac{E_\mathbf{p}+m}{2m}\left(1+\frac{E_\mathbf{p}-m}{E_\mathbf{p}+m}\right)=\frac{E_\mathbf{p}+m}{2m}\frac{2E_\mathbf{p}}{E_\mathbf{p}+m}=\frac{E_\mathbf{p}}{m}.\end{aligned} \tag{W.125}$$

[89] Remember $(\boldsymbol{\sigma}\mathbf{a})(\boldsymbol{\sigma}\mathbf{b}) = \mathbf{ab} + i\boldsymbol{\sigma}(\mathbf{a}\times\mathbf{b})$.

Finally we calculate $\bar{u}_r(-\mathbf{p})\gamma^0 w_{r'}(\mathbf{p}) = u_r^\dagger(-\mathbf{p})w_{r'}(\mathbf{p})$. We have $E_{-\mathbf{p}} = \sqrt{(-\mathbf{p})^2 + m^2} = E_\mathbf{p}$. It follows

$$\begin{aligned} u_r^\dagger(-\mathbf{p})w_{r'}(\mathbf{p}) &= -\left(\tfrac{E_\mathbf{p}+m}{2m}\right)^{1/2}\left(\tfrac{E_\mathbf{p}+m}{2m}\right)^{1/2}\left(\chi_r^\dagger \ \left(-\tfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E_\mathbf{p}+m}\chi_r\right)^\dagger\right)\begin{pmatrix}\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E_\mathbf{p}+m}i\sigma^2\chi_{r'} \\ i\sigma^2\chi_{r'}\end{pmatrix} = \\ &= -\tfrac{E_\mathbf{p}+m}{2m}\left[\chi_r^\dagger\tfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E_\mathbf{p}+m}i\sigma^2\chi_{r'} - \left(\tfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E_\mathbf{p}+m}\chi_r\right)^\dagger i\sigma^2\chi_{r'}\right] = \\ &= -\tfrac{E_\mathbf{p}+m}{2m}\left[\chi_r^\dagger\tfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E_\mathbf{p}+m}i\sigma^2\chi_{r'} - \chi_r^\dagger\left(\tfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E_\mathbf{p}+m}\right)i\sigma^2\chi_{r'}\right] = 0. \end{aligned} \tag{W.126}$$

7. Prove ([W.100](#)) for the discrete case.
Solution: With

$$H = i\int \mathrm{d}^3x\ \psi^\dagger(x)\,\partial_0\psi(x) \tag{W.127}$$

and

$$\begin{aligned} \psi(x) &= \sum\nolimits_{\mathbf{p},r}\sqrt{\tfrac{m}{VE_\mathbf{p}}}\left(b_r(\mathbf{p})\,u_r(\mathbf{p})\,e^{-ipx} + \mathrm{d}_r^\dagger(\mathbf{p})\,w_r(\mathbf{p})\,e^{ipx}\right) \\ \bar{\psi}(x) &= \sum\nolimits_{\mathbf{p},r}\sqrt{\tfrac{m}{VE_\mathbf{p}}}\left(\mathrm{d}_r(\mathbf{p})\,\bar{w}_r(\mathbf{p})\,e^{-ipx} + b_r^\dagger(\mathbf{p})\,\bar{u}_r(\mathbf{p})\,e^{ipx}\right) \end{aligned} \tag{W.128}$$

we arrive at

$$\begin{aligned} H &= i\int \mathrm{d}^3x\ \bar{\psi}(x)\,\gamma^0\partial_0\psi(x) = \\ &= i\sum\nolimits_{\mathbf{p},r,\mathbf{p}',r'}\sqrt{\tfrac{m}{VE_\mathbf{p}}}\sqrt{\tfrac{m}{VE_{\mathbf{p}'}}}\int \mathrm{d}^3x\ \left[\mathrm{d}_r(\mathbf{p})\,\bar{w}_r(\mathbf{p})\,e^{-ipx} + b_r^\dagger(\mathbf{p})\,\bar{u}_r(\mathbf{p})\,e^{ipx}\right]\cdot \\ &\cdot\gamma^0\partial_0\left[b_{r'}\left(\mathbf{p}'\right)u_{r'}\left(\mathbf{p}'\right)e^{-ip'x} + \mathrm{d}_{r'}^\dagger\left(\mathbf{p}'\right)w_{r'}\left(\mathbf{p}'\right)e^{ip'x}\right]. \end{aligned} \tag{W.129}$$

With

$$\partial_0 e^{ipx} = \partial_0 e^{i\left(E_\mathbf{p}t - \mathbf{p}x\right)} = iE_\mathbf{p}e^{ipx} \tag{W.130}$$

follows

$$\begin{aligned} H &= -\sum\nolimits_{\mathbf{p},r,\mathbf{p}',r'}\sqrt{\tfrac{m}{VE_\mathbf{p}}}\sqrt{\tfrac{m}{VE_{\mathbf{p}'}}}E_\mathbf{p}\cdot \\ &\cdot\int \mathrm{d}^3x\left[\begin{pmatrix} -\mathrm{d}_r(\mathbf{p})\,\bar{w}_r(p)\,\gamma^0 b_r'\left(\mathbf{p}'\right)u_{r'}\left(p'\right)e^{-ipx}e^{-ip'x} - b_r^\dagger(\mathbf{p})\,\bar{u}_r(p)\,\gamma^0 b_r'\left(\mathbf{p}'\right)u_{r'}\left(p'\right)e^{ipx}e^{-ip'x} + \\ +\mathrm{d}_r(\mathbf{p})\,\bar{w}_r(p)\,\gamma^0\mathrm{d}_{r'\mathbf{p}'}^\dagger w_{r'}\left(p'\right)e^{-ipx}e^{ip'x} + b_r^\dagger(\mathbf{p})\,\bar{u}_r(p)\,\gamma^0\mathrm{d}_{r'}^\dagger\left(\mathbf{p}'\right)w_{r'}\left(p'\right)e^{ipx}e^{ip'x}\end{pmatrix}\right]. \end{aligned} \tag{W.131}$$

Performing the x-integration gives (see Appendix T, Vol. 1)

$$H = -V\sum_{\mathbf{p},r,\mathbf{p}',r'}\sqrt{\frac{m}{VE_{\mathbf{p}}}}\sqrt{\frac{m}{VE_{\mathbf{p}'}}}E_{\mathbf{p}}\left[\left(\begin{array}{c} -\mathrm{d}_r(\mathbf{p})\,\bar{w}_r(\mathbf{p})\,\gamma^0 b'_r(\mathbf{p}')\,u_{r'}(\mathbf{p}')\,e^{-ip_0x^0}e^{-ip'_0x^0}\delta_{\mathbf{p}+\mathbf{p}',\mathbf{0}}- \\ -b_r^\dagger(\mathbf{p})\,\bar{u}_r(\mathbf{p})\,\gamma^0 b'_r(\mathbf{p}')\,u_{r'}(\mathbf{p}')\,e^{ip_0x^0}e^{-ip'_0x^0}\delta_{\mathbf{p}-\mathbf{p}',\mathbf{0}}+ \\ +\mathrm{d}_r(\mathbf{p})\,\bar{w}_r(\mathbf{p})\,\gamma^0\mathrm{d}_{r'}^\dagger(\mathbf{p}')\,w_{r'}(\mathbf{p}')\,e^{-ip_0x^0}e^{ip'_0x^0}\delta_{\mathbf{p}-\mathbf{p}',\mathbf{0}}+ \\ +b_r^\dagger(\mathbf{p})\,\bar{u}_r(\mathbf{p})\,\gamma^0\mathrm{d}_{r'}^\dagger(\mathbf{p}')\,w_{r'}(\mathbf{p}')\,e^{ip_0x^0}e^{ip'_0x^0}\delta_{\mathbf{p}+\mathbf{p}',\mathbf{0}} \end{array}\right)\right] =$$

$$= -\sum_{\mathbf{p},r,\mathbf{p}',r'}\sqrt{\frac{m}{E_{\mathbf{p}}}}\sqrt{\frac{m}{E_{\mathbf{p}'}}}E_{\mathbf{p}}\left[\left(\begin{array}{c} \delta_{\mathbf{p}+\mathbf{p}',\mathbf{0}}\left[\begin{array}{c} -\mathrm{d}_r(\mathbf{p})\,\bar{w}_r(\mathbf{p})\,\gamma^0 b'_r(\mathbf{p}')\,u_{r'}(\mathbf{p}')\,e^{-ip_0x^0}e^{-ip'_0x^0}+ \\ +b_r^\dagger(\mathbf{p})\,\bar{u}_r(\mathbf{p})\,\gamma^0\mathrm{d}_{r'}^\dagger(\mathbf{p}')\,w_{r'}(\mathbf{p}')\,e^{ip_0x^0}e^{ip'_0x^0} \end{array}\right]+ \\ +\delta_{\mathbf{p}-\mathbf{p}',\mathbf{0}}\left[\begin{array}{c} -b_r^\dagger(\mathbf{p})\,\bar{u}_r(\mathbf{p})\,\gamma^0 b'_r(\mathbf{p}')\,u_{r'}(\mathbf{p}')\,e^{ip_0x^0}e^{-ip'_0x^0}+ \\ +\mathrm{d}_r(\mathbf{p})\,\bar{w}_r(\mathbf{p})\,\gamma^0\mathrm{d}_{r'}^\dagger(\mathbf{p}')\,w_{r'}(\mathbf{p}')\,e^{-ip_0x^0}e^{ip'_0x^0} \end{array}\right] \end{array}\right)\right].$$
(W.132)

Due to the Kronecker deltas we have $E_{\mathbf{p}'} = \sqrt{\mathbf{p}'^2 + m^2} = \sqrt{(\pm\mathbf{p})^2 + m^2} = E_{\mathbf{p}}$, i.e., $p_0 = E_{\mathbf{p}} = p'_0 = E_{\mathbf{p}'}$ It follows

$$H = \left(\begin{array}{c} -m\sum_{\mathbf{p},r,r'}\left[\begin{array}{c} -\mathrm{d}_r(\mathbf{p})\,\bar{w}_r(\mathbf{p})\,\gamma^0 b'_r(-\mathbf{p})\,u_{r'}(-\mathbf{p})\,e^{-2ip_0x^0}+ \\ +b_r^\dagger(\mathbf{p})\,\bar{u}_r(\mathbf{p})\,\gamma^0\mathrm{d}_{r'}^\dagger(-\mathbf{p})\,w_{r'}(-\mathbf{p})\,e^{2ip_0x^0} \end{array}\right]- \\ -m\sum_{\mathbf{p},r,r'}\left[\begin{array}{c} -b_r^\dagger(\mathbf{p})\,\bar{u}_r(\mathbf{p})\,\gamma^0 b'_r(\mathbf{p})\,u_{r'}(\mathbf{p})+ \\ +\mathrm{d}_r(\mathbf{p})\,\bar{w}_r(\mathbf{p})\,\gamma^0\mathrm{d}_{r'}^\dagger(\mathbf{p})\,w_{r'}(\mathbf{p}) \end{array}\right] \end{array}\right). \tag{W.133}$$

Due to (W.99), i.e.,

$$\begin{array}{c} \bar{u}_r(\mathbf{p})\,\gamma^0 u_{r'}(\mathbf{p}) = \frac{E_{\mathbf{p}}}{m}\delta_{rr'}\ ;\ \bar{u}_r(-\mathbf{p})\,\gamma^0 w_{r'}(\mathbf{p}) = 0 \\ \bar{w}_r(\mathbf{p})\,\gamma^0 w_{r'}(\mathbf{p}) = \frac{E_{\mathbf{p}}}{m}\delta_{rr'}\ ;\ \bar{w}_r(-\mathbf{p})\,\gamma^0 u_{r'}(\mathbf{p}) = 0 \end{array} \tag{W.134}$$

follows finally

$$H = -m\sum_{\mathbf{p},r,r'}\left[-b_r^\dagger(\mathbf{p})\,\bar{u}_r(\mathbf{p})\,\gamma^0 b_{r'}(\mathbf{p}')\,u_{r'}(\mathbf{p}) + \mathrm{d}_r(\mathbf{p})\,\bar{w}_r(\mathbf{p})\,\gamma^0\mathrm{d}_{r'}^\dagger(\mathbf{p})\,w_{r'}(\mathbf{p})\right] =$$

$$= -m\sum_{\mathbf{p},r}\frac{E_{\mathbf{p}}}{m}\left[-b_r^\dagger(\mathbf{p})\,b_r(\mathbf{p}) + \mathrm{d}_r(\mathbf{p})\,\mathrm{d}_r^\dagger(\mathbf{p})\right] = \sum_{\mathbf{p},r}E_{\mathbf{p}}\left[b_r^\dagger(\mathbf{p})\,b_r(\mathbf{p}) - \mathrm{d}_r(\mathbf{p})\,\mathrm{d}_r^\dagger(\mathbf{p})\right]. \tag{W.135}$$

8. Prove (W.100) for the continuous case.
Solution: We have

$$H = i\int \mathrm{d}^3x\ \psi^\dagger\partial_0\psi \tag{W.136}$$

and

$$\begin{array}{c} \psi = \sum_r \int \mathrm{d}^3p\,\sqrt{\frac{m}{(2\pi)^3E_{\mathbf{p}}}}\left(b_r(\mathbf{p})\,u_r(\mathbf{p})\,e^{-ipx} + \mathrm{d}_r^\dagger(\mathbf{p})\,w_r(\mathbf{p})\,e^{ipx}\right) \\ \psi^\dagger = \sum_r \int \mathrm{d}^3p\,\sqrt{\frac{m}{(2\pi)^3E_{\mathbf{p}}}}\left(\mathrm{d}_r(\mathbf{p})\,w_r^\dagger(\mathbf{p})\,e^{-ipx} + b_r^\dagger(\mathbf{p})\,u_r^\dagger(\mathbf{p})\,e^{ipx}\right). \end{array} \tag{W.137}$$

With $\partial_0 e^{ipx} = ip_0e^{ipx}$, $\partial_0\psi$ is given by

$$\partial_0\psi = i\sum_r \sqrt{\frac{m}{(2\pi)^3}} \int \frac{d^3p}{\sqrt{E_\mathbf{p}}} p_0 \left[-b_r(\mathbf{p})\, u_r(\mathbf{p})\, e^{-ipx} + d_r^\dagger(\mathbf{p})\, w_r(\mathbf{p})\, e^{ipx}\right]. \tag{W.138}$$

It follows

$$H = -\int \frac{d^3p}{\sqrt{E_\mathbf{p}}} \int \frac{d^3p'}{\sqrt{E_{\mathbf{p}'}}} p'_0 \sum_{r.r'} \sqrt{\frac{m}{(2\pi)^3}} \sqrt{\frac{m}{(2\pi)^3}} \int d^3x$$
$$\left[d_r(\mathbf{p})\, w_r^\dagger(\mathbf{p})\, e^{-ipx} + b_r^\dagger(\mathbf{p})\, u_r^\dagger(\mathbf{p})\, e^{ipx}\right]\left[-b_{r'}(\mathbf{p}')\, u_{r'}(\mathbf{p}')\, e^{-ip'x} + d_{r'}^\dagger(\mathbf{p}')\, w_{r'}(\mathbf{p}')\, e^{ip'x}\right]. \tag{W.139}$$

We consider the x-integration:

$$\begin{aligned} I &= \int d^3x \left[d_r(\mathbf{p})\, w_r^\dagger(\mathbf{p})\, e^{-ipx} + b_r^\dagger(\mathbf{p})\, u_r^\dagger(\mathbf{p})\, e^{ipx}\right]\left[-b_{r'}(\mathbf{p}')\, u_{r'}(\mathbf{p}')\, e^{-ip'x} + d_{r'}^\dagger(\mathbf{p}')\, w_{r'}(\mathbf{p}')\, e^{ip'x}\right] = \\ &= \int d^3x \left[\begin{array}{c} \left[-d_r(\mathbf{p})\, w_r^\dagger(\mathbf{p})\, b_{r'}(\mathbf{p}')\, u_{r'}(\mathbf{p}')\, e^{-i(p'+p)x} + d_r(\mathbf{p})\, w_r^\dagger(\mathbf{p})\, d_{r'}^\dagger(\mathbf{p}')\, w_{r'}(\mathbf{p}')\, e^{i(p'-p)x}\right] + \\ + \left[-b_r^\dagger(\mathbf{p})\, u_r^\dagger(\mathbf{p})\, b_{r'}(\mathbf{p}')\, u_{r'}(\mathbf{p}')\, e^{-i(p'-p)x} + b_r^\dagger(\mathbf{p})\, u_r^\dagger(\mathbf{p})\, d_{r'}^\dagger(\mathbf{p}')\, w_{r'}(\mathbf{p}')\, e^{i(p'+p)x}\right] \end{array}\right]. \end{aligned} \tag{W.140}$$

The x-integration yields

$$\int d^3x\, e^{ikx} = e^{ik_0x^0} \int d^3x\, e^{-i\mathbf{kx}} = e^{ik_0x^0} (2\pi)^3 \delta^3(\mathbf{k}). \tag{W.141}$$

It follows

$$I = (2\pi)^3 \left[\begin{array}{c} \left[\begin{array}{c} -d_r(\mathbf{p})\, w_r^\dagger(\mathbf{p})\, b_{r'}(\mathbf{p}')\, u_{r'}(\mathbf{p}')\, e^{-i(p'_0+p_0)x^0} \delta^3(\mathbf{p}'+\mathbf{p}) + \\ +d_r(\mathbf{p})\, w_r^\dagger(\mathbf{p})\, d_{r'}^\dagger(\mathbf{p}')\, w_{r'}(\mathbf{p}')\, e^{i(p'_0-p_0)x^0} \delta^3(\mathbf{p}'-\mathbf{p}) \end{array}\right] + \\ + \left[\begin{array}{c} -b_r^\dagger(\mathbf{p})\, u_r^\dagger(\mathbf{p})\, b_{r'}(\mathbf{p}')\, u_{r'}(\mathbf{p}')\, e^{-i(p'_0-p_0)x^0} \delta^3(\mathbf{p}'-\mathbf{p}) + \\ +b_r^\dagger(\mathbf{p})\, u_r^\dagger(\mathbf{p})\, d_{r'}^\dagger(\mathbf{p}')\, w_{r'}(\mathbf{p}')\, e^{i(p'_0+p_0)x^0} \delta^3(\mathbf{p}'+\mathbf{p}) \end{array}\right] \end{array}\right]. \tag{W.142}$$

The Kronecker functions give $E_{\mathbf{p}'} = \sqrt{\mathbf{p}'^2 + m^2} = \sqrt{(\pm\mathbf{p})^2 + m^2} = E_\mathbf{p}$, i.e., $p_0 = E_\mathbf{p} = p'_0 = E_{\mathbf{p}'}$. Thus, we arrive at

$$I = (2\pi)^3 \left[\begin{array}{c} \delta^3(\mathbf{p}'-\mathbf{p}) \left[\begin{array}{c} d_r(\mathbf{p})\, d_{r'}^\dagger(\mathbf{p})\, w_r^\dagger(\mathbf{p})\, w_{r'}(\mathbf{p}) - \\ -b_r^\dagger(\mathbf{p})\, b_{r'}(\mathbf{p})\, u_r^\dagger(\mathbf{p})\, u_{r'}(\mathbf{p}) \end{array}\right] + \\ +\delta^3(\mathbf{p}'+\mathbf{p}) \left[\begin{array}{c} b_r^\dagger(\mathbf{p})\, d_{r'}^\dagger(-\mathbf{p})\, u_r^\dagger(\mathbf{p})\, w_{r'}(-\mathbf{p})\, e^{2ip_0x^0} - \\ -d_r(\mathbf{p})\, b_{r'}(-\mathbf{p})\, w_r^\dagger(\mathbf{p})\, u_{r'}(-\mathbf{p})\, e^{-2ip_0x^0} \end{array}\right] \end{array}\right]. \tag{W.143}$$

Inserting (W.99), i.e.,

$$u_r^\dagger(\mathbf{p})\, w_{r'}(-\mathbf{p}) = w_r^\dagger(\mathbf{p})\, u_{r'}(-\mathbf{p}) = 0 \tag{W.144}$$

yields

$$I = (2\pi)^3 \delta^3(\mathbf{p}'-\mathbf{p}) \left[\begin{array}{c} d_r(\mathbf{p})\, d_{r'}^\dagger(\mathbf{p})\, w_r^\dagger(\mathbf{p})\, w_{r'}(\mathbf{p}) - \\ -b_r^\dagger(\mathbf{p})\, b_{r'}(\mathbf{p})\, u_r^\dagger(\mathbf{p})\, u_{r'}(\mathbf{p}) \end{array}\right]. \tag{W.145}$$

In addition, we use (W.98), i.e.,

$$u_r^\dagger(\mathbf{p})\,u_{r'}(\mathbf{p}) = w_r^\dagger(\mathbf{p})\,w_{r'}(\mathbf{p}) = \frac{E_\mathbf{p}}{m}\delta_{rr'} \tag{W.146}$$

and obtain

$$I = (2\pi)^3\,\delta^3\left(\mathbf{p}' - \mathbf{p}\right)\frac{E_\mathbf{p}}{m}\delta_{rr'}\left[\mathrm{d}_r(\mathbf{p})\,\mathrm{d}_{r'}^\dagger(\mathbf{p}) - b_r^\dagger(\mathbf{p})\,b_{r'}(\mathbf{p})\right]. \tag{W.147}$$

We insert this result for I and obtain

$$\begin{aligned} H - \int \frac{\mathrm{d}^3 p}{\sqrt{E_\mathbf{p}}} \int \frac{\mathrm{d}^3 p'}{\sqrt{E_{\mathbf{p}'}}} p_0' \textstyle\sum_{r.r'} \sqrt{\frac{m}{(2\pi)^3}}\sqrt{\frac{m}{(2\pi)^3}} \cdot I = \\ = \int \frac{\mathrm{d}^3 p}{\sqrt{E_\mathbf{p}}} \int \frac{\mathrm{d}^3 p'}{\sqrt{E_{\mathbf{p}'}}} p_0' \textstyle\sum_r \delta^3\left(\mathbf{p}' - \mathbf{p}\right) E_\mathbf{p}\left[b_r^\dagger(\mathbf{p})\,b_r(\mathbf{p}) - \mathrm{d}_r(\mathbf{p})\,\mathrm{d}_r^\dagger(\mathbf{p})\right] = \\ = \int \mathrm{d}^3 p\, E_\mathbf{p} \textstyle\sum_r \left[b_r^\dagger(\mathbf{p})\,b_r(\mathbf{p}) - \mathrm{d}_r(\mathbf{p})\,\mathrm{d}_r^\dagger(\mathbf{p})\right]. \end{aligned} \tag{W.148}$$

W.6 Quantization of Free Fields, Photons

Essential issues are already discussed in the section 'Toy example' and in Appendix T, Vol. 1. We want to add some more results here in order to establish an uniform formalism.

A suitable Lagrangian density $\mathcal{L}$ is given by

$$\mathcal{L} = -\frac{1}{2}\left(\partial_\nu A_\mu\right)\left(\partial^\nu A^\mu\right). \tag{W.149}$$

Note that we have four fields, namely A^μ, $\mu = 0, \ldots, 3$. The corresponding conjugated fields are given by

$$\pi_\mu = \frac{\delta\mathcal{L}}{\delta\left(\partial_0 A^\mu\right)} = -\partial_0 A_\mu. \tag{W.150}$$

W.6.1 Determination of $\mathcal{H}$

In the next step, we calculate the Hamiltonian $\mathcal{H}$. We have

$$\begin{aligned} \mathcal{H} = \pi_\mu \partial^0 A^\mu - \mathcal{L} = -\left(\partial_0 A_\mu\right)\left(\partial_0 A^\mu\right) - \mathcal{L} = \\ = -\left(\partial_0 A_\mu\right)\left(\partial^0 A^\mu\right) + \tfrac{1}{2}\left(\partial_\nu A_\mu\right)\left(\partial^\nu A^\mu\right) = \\ = -\left(\partial_0 A_\mu\right)\left(\partial^0 A^\mu\right) + \tfrac{1}{2}\left[\left(\partial_0 A_\mu\right)\left(\partial^0 A^\mu\right) + \left(\partial_k A_\mu\right)\left(\partial^k A^\mu\right)\right] = \\ = -\tfrac{1}{2}\left(\partial_0 A_\mu\right)\left(\partial^0 A^\mu\right) + \tfrac{1}{2}\left(\partial_k A_\mu\right)\left(\partial^k A^\mu\right). \end{aligned} \tag{W.151}$$

With $\partial_k = -\partial^k$, this may be written as $\mathcal{H} = -\frac{1}{2}\left(\partial_0 A_\mu\right)\left(\partial^0 A^\mu\right) - \frac{1}{2}\left(\partial^k A_\mu\right)\left(\partial^k A^\mu\right)$ or

$$\mathcal{H} = -\frac{1}{2}\left(\partial^\nu A_\mu\right)\left(\partial^\nu A^\mu\right). \tag{W.152}$$

The discrete and continuous solutions in the source-free case read (see Appendix T, Vol. 1)[90]

$$\begin{aligned} A^\mu(x) = \textstyle\sum_{\mathbf{k},r} \sqrt{\frac{1}{2V\omega_\mathbf{k}}} \varepsilon_r^\mu(\mathbf{k})\left[\alpha_r(\mathbf{k})\, e^{-ikx} + \alpha_r^\dagger(\mathbf{k})\, e^{ikx}\right] \\ A^\mu(x) = \textstyle\sum_r \int \frac{\mathrm{d}^3 k}{\sqrt{2(2\pi)^3\omega_\mathbf{k}}}\, \varepsilon_r^\mu(\mathbf{k})\left[\alpha_r(\mathbf{k})\, e^{-ikx} + \alpha_r^\dagger(\mathbf{k})\, e^{ikx}\right]. \end{aligned} \tag{W.153}$$

If we choose the Coulomb gauge, then the polarization vectors have the properties

$$\mathbf{k} \cdot \boldsymbol{\varepsilon}_r(\mathbf{k}) = 0\ ;\ \varepsilon_r^0(\mathbf{k}) = 0\ ;\ \boldsymbol{\varepsilon}_r(\mathbf{k}) \cdot \boldsymbol{\varepsilon}_s(\mathbf{k}) = \delta_{rs}. \tag{W.154}$$

We now interpret the amplitudes $\alpha_r(\mathbf{k})$ as operators which makes the field $A^\mu(x)$ a field operator. Instead of performing the canonical quantization procedure, we calculate the energy H in terms of $\alpha_r(\mathbf{k})$ and $\alpha_r^\dagger(\mathbf{k})$ and discuss then which commutation relations are suitable. For the sake of variety, we start from the continuous solution.

We need for the (lengthy) calculation the expressions $\partial^\nu A_\mu$ and $\partial^\nu A^\mu$. With

$$\partial^\nu e^{ikx} = ik_\nu e^{ikx}\ ;\ k_\nu = \left(k^0, -\mathbf{k}\right)\ ;\ \varepsilon_{\nu r}(\mathbf{k}) = g_{\nu\mu}\varepsilon_r^\mu(\mathbf{k}) \tag{W.155}$$

follows

$$\begin{aligned} \partial^\nu A_\mu = \textstyle\sum_r \int \frac{\mathrm{d}^3 k}{\sqrt{2(2\pi)^3\omega_\mathbf{k}}}\, \varepsilon_{\mu r}(\mathbf{k})\, ik_\nu\left[-\alpha_r(\mathbf{k})\, e^{-ikx} + \alpha_r^\dagger(\mathbf{k})\, e^{ikx}\right] \\ \partial^\nu A^\mu = \textstyle\sum_r \int \frac{\mathrm{d}^3 k}{\sqrt{2(2\pi)^3\omega_\mathbf{k}}}\, \varepsilon_r^\mu(\mathbf{k})\, ik_\nu\left[-\alpha_r(\mathbf{k})\, e^{-ikx} + \alpha_r^\dagger(\mathbf{k})\, e^{ikx}\right]. \end{aligned} \tag{W.156}$$

Thus, the Hamiltonian is given by

$$\begin{aligned} -2\mathcal{H} = \left(\partial^\nu A_\mu\right)\left(\partial^\nu A^\mu\right) = \\ = \textstyle\sum_r \int \frac{\mathrm{d}^3 k}{\sqrt{2(2\pi)^3\omega_\mathbf{k}}}\, \varepsilon_{\mu r}(\mathbf{k})\, ik_\nu\left[-\alpha_r(\mathbf{k})\, e^{-ikx} + \alpha_r^\dagger(\mathbf{k})\, e^{ikx}\right] \cdot \\ \cdot \textstyle\sum_{r'} \int \frac{\mathrm{d}^3 k'}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}'}}}\, \varepsilon_{r'}^\mu\left(\mathbf{k}'\right) ik'_\nu\left[-\alpha_{r'}\left(\mathbf{k}'\right) e^{-ik'x} + \alpha_{r'}^\dagger\left(\mathbf{k}'\right) e^{ikx'}\right]. \end{aligned} \tag{W.157}$$

---

[90] Note $\omega_\mathbf{k} = k^0$ and $\omega_\mathbf{k}^2 = (k^0)^2 = \mathbf{k}^2$.

W.6.2 Determination of H

For the energy follows

$$\begin{gathered}-2H=-2\int \mathrm{d}^3x\ \mathcal{H}(x)=\int \mathrm{d}^3x\ \left(\partial^\nu A_\mu\right)\left(\partial^\nu A^\mu\right)= \\ =\begin{array}{c}\int \mathrm{d}^3x\ \sum_{rr'}\int \frac{\mathrm{d}^3k}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}}\int \frac{\mathrm{d}^3k'}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}'}}}\varepsilon_{\mu r}(\mathbf{k})\,\varepsilon^\mu_{r'}\left(\mathbf{k}'\right)(ik_\nu)\left(ik'_\nu\right)\cdot \\ \cdot\left[-\alpha_r(\mathbf{k})\,e^{-ikx}+\alpha_r^\dagger(\mathbf{k})\,e^{ikx}\right]\cdot\left[-\alpha_{r'}\left(\mathbf{k}'\right)e^{-ik'x}+\alpha_{r'}^\dagger\left(\mathbf{k}'\right)e^{ikx'}\right]=\end{array} \\ =\sum_{rr'}\int \frac{\mathrm{d}^3k}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}}\int \frac{\mathrm{d}^3k'}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}'}}}\varepsilon_{\mu r}(\mathbf{k})\,\varepsilon^\mu_{r'}\left(\mathbf{k}'\right)(ik_\nu)\left(ik'_\nu\right)\cdot I\left(r,r',\mathbf{k},\mathbf{k}'\right)\end{gathered} \tag{W.158}$$

with

$$I\left(r,r',\mathbf{k},\mathbf{k}'\right)=\int \mathrm{d}^3x\left[-\alpha_r(\mathbf{k})\,e^{-ikx}+\alpha_r^\dagger(\mathbf{k})\,e^{ikx}\right]\cdot\left[-\alpha_{r'}\left(\mathbf{k}'\right)e^{-ik'x}+\alpha_{r'}^\dagger\left(\mathbf{k}'\right)e^{ikx'}\right]. \tag{W.159}$$

With

$$\int \mathrm{d}^3x\ e^{ipx}=e^{ip_0x_0}\int \mathrm{d}^3x\ e^{-i\mathbf{px}}=e^{ip_0x_0}\,(2\pi)^3\,\delta(\mathbf{p}) \tag{W.160}$$

follows

$$\begin{gathered}I\left(r,r',\mathbf{k},\mathbf{k}'\right)=\int \mathrm{d}^3x\left[-\alpha_r(\mathbf{k})\,e^{-ikx}+\alpha_r^\dagger(\mathbf{k})\,e^{ikx}\right]\cdot\left[-\alpha_{r'}\left(\mathbf{k}'\right)e^{-ik'x}+\alpha_{r'}^\dagger\left(\mathbf{k}'\right)e^{ikx'}\right]= \\ =\int \mathrm{d}^3x\ \cdot\left[\begin{array}{c}\alpha_r(\mathbf{k})\,\alpha_{r'}\left(\mathbf{k}'\right)e^{-i\left(k+k'\right)x}-\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}\left(\mathbf{k}'\right)e^{i\left(k-k'\right)x}- \\ -\alpha_r(\mathbf{k})\,\alpha_{r'}^\dagger\left(\mathbf{k}'\right)e^{-i\left(k-k'\right)x}+\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}^\dagger\left(\mathbf{k}'\right)e^{i\left(k+k'\right)x}\end{array}\right]= \\ =(2\pi)^3\cdot\left[\begin{array}{c}\left\{\alpha_r(\mathbf{k})\,\alpha_{r'}\left(\mathbf{k}'\right)e^{-i\left(k_0+k'_0\right)x_0}+\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}^\dagger\left(\mathbf{k}'\right)e^{i\left(k_0+k'_0\right)x_0}\right\}\delta\left(\mathbf{k}+\mathbf{k}'\right)- \\ -\left\{\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}\left(\mathbf{k}'\right)e^{i\left(k_0-k'_0\right)x_0}+\alpha_r(\mathbf{k})\,\alpha_{r'}^\dagger\left(\mathbf{k}'\right)e^{-i\left(k_0-k'_0\right)x_0}\right\}\delta\left(\mathbf{k}-\mathbf{k}'\right)\end{array}\right].\end{gathered} \tag{W.161}$$

Due to the delta functions, only the following terms survive:

$$\delta\left(\mathbf{k}+\mathbf{k}'\right)\to\mathbf{k}'=-\mathbf{k}\ ;\ k'_0=k_0\ ;\ \delta\left(\mathbf{k}-\mathbf{k}'\right)\to\mathbf{k}'=\mathbf{k}\ ;\ k'_0=k_0. \tag{W.162}$$

This yields

$$I\left(r,r',\mathbf{k},\mathbf{k}'\right)=(2\pi)^3\cdot\left[\begin{array}{c}\left\{\alpha_r(\mathbf{k})\,\alpha_{r'}(-\mathbf{k})\,e^{-2ik_0x}+\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}^\dagger(-\mathbf{k})\,e^{2ik_0}\right\}\delta\left(\mathbf{k}+\mathbf{k}'\right)- \\ -\left\{\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}(\mathbf{k})+\alpha_r(\mathbf{k})\,\alpha_{r'}^\dagger(\mathbf{k})\right\}\delta\left(\mathbf{k}-\mathbf{k}'\right)\end{array}\right]. \tag{W.163}$$

Thus, we arrive at

$$
\begin{aligned}
-2H &= -2\int \mathrm{d}^3x\ \mathcal{H}(x) = \\
&= \sum_{rr'} \int \frac{\mathrm{d}^3k}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}} \int \frac{\mathrm{d}^3k'}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}'}}}\ \varepsilon_{\mu r}(\mathbf{k})\,\varepsilon^{\mu}_{r'}(\mathbf{k}')\,(ik_\nu)\left(ik'_\nu\right)(2\pi)^3 \cdot \\
&\cdot \left[\begin{array}{c} \left\{\alpha_r(\mathbf{k})\,\alpha_{r'}(-\mathbf{k})\,e^{-2ik_0x} + \alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}^\dagger(-\mathbf{k})\,e^{2ik_0}\right\}\delta\left(\mathbf{k}+\mathbf{k}'\right) - \\ -\left\{\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}(\mathbf{k}) + \alpha_r(\mathbf{k})\,\alpha_{r'}^\dagger(\mathbf{k})\right\}\delta\left(\mathbf{k}-\mathbf{k}'\right) \end{array}\right] = \\
&= \sum_{rr'} \int \frac{\mathrm{d}^3k}{2\omega_{\mathbf{k}}} \int \frac{\mathrm{d}^3k'}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}'}}} \left[\begin{array}{c} \varepsilon_{\mu r}(\mathbf{k})\,\varepsilon^{\mu}_{r'}(-\mathbf{k})\,(ik_\nu)\left(ik'_\nu\right)\left\{\begin{array}{c}\alpha_r(\mathbf{k})\,\alpha_{r'}(-\mathbf{k})\,e^{-2ik_0x} + \\ +\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}^\dagger(-\mathbf{k})\,e^{2ik_0}\end{array}\right\}\delta\left(\mathbf{k}+\mathbf{k}'\right) - \\ -\varepsilon_{\mu r}(\mathbf{k})\,\varepsilon^{\mu}_{r'}(\mathbf{k})\,(ik_\nu)\left(ik'_\nu\right)\left\{\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}(\mathbf{k}) + \alpha_r(\mathbf{k})\,\alpha_{r'}^\dagger(\mathbf{k})\right\}\delta\left(\mathbf{k}-\mathbf{k}'\right) \end{array}\right].
\end{aligned}
\tag{W.164}
$$

Now we consider the terms $(ik_\nu)\left(ik'_\nu\right)\delta\left(\mathbf{k}+\mathbf{k}'\right)$ and $(ik_\nu)\left(ik'_\nu\right)\delta\left(\mathbf{k}-\mathbf{k}'\right)$. We have (remember $k_0^2 = \mathbf{k}^2$)

$$
\begin{aligned}
(ik_\nu)\left(ik'_\nu\right)\delta\left(\mathbf{k}+\mathbf{k}'\right) &= -\left(k_0^2 + \mathbf{k}\mathbf{k}'\right)\delta\left(\mathbf{k}+\mathbf{k}'\right) = -\left(k_0^2 - \mathbf{k}\mathbf{k}\right)\delta\left(\mathbf{k}+\mathbf{k}'\right) = 0 \\
(ik_\nu)\left(ik'_\nu\right)\delta\left(\mathbf{k}-\mathbf{k}'\right) &= -\left(k_0^2 + \mathbf{k}\mathbf{k}\right)\delta\left(\mathbf{k}-\mathbf{k}'\right) = -2\omega_{\mathbf{k}}^2\delta\left(\mathbf{k}-\mathbf{k}'\right).
\end{aligned}
\tag{W.165}
$$

This yields

$$
\begin{aligned}
-2H &= \sum_{rr'} \int \frac{\mathrm{d}^3k}{2\omega_{\mathbf{k}}} \left[-\varepsilon_{\mu r}(\mathbf{k})\,\varepsilon^{\mu}_{r'}(\mathbf{k})\left(-2\omega_{\mathbf{k}}^2\right)\left\{\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}(\mathbf{k}) + \alpha_r(\mathbf{k})\,\alpha_{r'}^\dagger(\mathbf{k})\right\}\right] = \\
&= -\sum_{rr'} \int \mathrm{d}^3k\ \omega_{\mathbf{k}}\left[\left(-\varepsilon_{\mu r}(\mathbf{k})\,\varepsilon^{\mu}_{r'}(\mathbf{k})\right)\left\{\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}(\mathbf{k}) + \alpha_r(\mathbf{k})\,\alpha_{r'}^\dagger(\mathbf{k})\right\}\right].
\end{aligned}
\tag{W.166}
$$

The product of the polarization vectors gives

$$
-\varepsilon_{\mu r}(\mathbf{k})\,\varepsilon^{\mu}_{r'}(\mathbf{k}) = -\varepsilon_{0r}(\mathbf{k})\,\varepsilon^{0}_{r'}(\mathbf{k}) - \left(-\boldsymbol{\varepsilon}_r(\mathbf{k})\,\boldsymbol{\varepsilon}_{r'}(\mathbf{k})\right) = -\varepsilon_{0r}(\mathbf{k})\,\varepsilon^{0}_{r'}(\mathbf{k}) + \boldsymbol{\varepsilon}_r(\mathbf{k})\,\boldsymbol{\varepsilon}_{r'}(\mathbf{k}).
\tag{W.167}
$$

In the source-free case, the Coulomb gauge is convenient. With (W.154) follows

$$
-\varepsilon_{\mu r}(\mathbf{k})\,\varepsilon^{\mu}_{r'}(\mathbf{k}) = \boldsymbol{\varepsilon}_r(\mathbf{k})\,\boldsymbol{\varepsilon}_{r'}(\mathbf{k}) = \delta_{rr'}
\tag{W.168}
$$

and we obtain

$$
\begin{aligned}
-2H &= -\sum_{rr'} \int \mathrm{d}^3k\ \omega_{\mathbf{k}}\delta_{rr'}\left\{\alpha_r^\dagger(\mathbf{k})\,\alpha_{r'}(\mathbf{k}) + \alpha_r(\mathbf{k})\,\alpha_{r'}^\dagger(\mathbf{k})\right\} \\
&= -\sum_{r} \int \mathrm{d}^3k\ \omega_{\mathbf{k}}\left\{\alpha_r^\dagger(\mathbf{k})\,\alpha_{r}(\mathbf{k}) + \alpha_r(\mathbf{k})\,\alpha_{r}^\dagger(\mathbf{k})\right\}.
\end{aligned}
\tag{W.169}
$$

Thus, the final result reads

$$
H = \sum_r \int \mathrm{d}^3k\ \omega_{\mathbf{k}} \frac{\alpha_r^\dagger(\mathbf{k})\,\alpha_r(\mathbf{k}) + \alpha_r(\mathbf{k})\,\alpha_r^\dagger(\mathbf{k})}{2}.
\tag{W.170}
$$

and in the discrete case

$$
H = \sum_{\mathbf{k},r} \omega_{\mathbf{k}}\left[\frac{\alpha_r^\dagger(\mathbf{k})\,\alpha_r(\mathbf{k}) + \alpha_r(\mathbf{k})\,\alpha_r^\dagger(\mathbf{k})}{2}\right].
\tag{W.171}
$$

We see that we have reproduced the result found in the section ‘Toy example’.

The commutation rules for the photon field are

$$\left[\alpha_r(\mathbf{k}), \alpha^\dagger_{r'}\left(\mathbf{k}'\right)\right] = \delta_{rr'}\delta\left(\mathbf{k}, \mathbf{k}'\right). \tag{W.172}$$

This may explicitly shown by the (quite lengthy) procedure of canonical quantization. But the result seems plausible anyway, apart from the fact, that we have deduced it already in a previous section. First, a photon is a boson; so we expect a commutator, not an anticommutator. Second, the state of a photon is completely determined by its momentum and its polarization. The further results can now be formulated parallel to those of the other fields. For instance, the number operator us given by $N_{\mathbf{k}r} = a^\dagger_r(\mathbf{k})\, a_r(\mathbf{k})$ and so on.

By means of the commutation rules, we can write

$$H = \sum_{\mathbf{k},r} \omega_{\mathbf{k}} \left[\alpha^\dagger_r(\mathbf{k})\, \alpha_r(\mathbf{k}) + \frac{1}{2}\right] \tag{W.173}$$

We see that also in this case we have an infinite vacuum energy.

W.6.3 Exercises and Solutions

1. Prove (W.150).
   Solution: We have

$$\begin{aligned} \pi_\mu = \tfrac{\delta\mathcal{L}}{\delta(\partial_0 A^\mu)} &= -\tfrac{1}{2}\tfrac{\partial}{\partial(\partial_0 A^\mu)}\left(\partial^\kappa A^\nu\right)\left(\partial_\kappa A_\nu\right) = \\ &= -\tfrac{1}{2}\tfrac{\partial}{\partial(\partial_0 A^\mu)}\left[\left(\partial^0 A^\nu\right)\left(\partial_0 A_\nu\right) + \left(\partial^k A^\mu\right)\left(\partial_k A_\mu\right)\right]. \end{aligned} \tag{W.174}$$

   The second summand contains no derivatives with respect to ∂_0 or ∂^0 and does not contribute to the result. It follows

$$\begin{aligned} \pi_\mu &= -\tfrac{1}{2}\tfrac{\partial}{\partial\left(\partial_0 A_\mu\right)}\left(\partial_0 A_\nu\right)\left(\partial^0 A^\nu\right) = \\ &= -\tfrac{1}{2}\left[\left(\tfrac{\partial}{\partial(\partial_0 A^\mu)}\left(\partial^0 A^\nu\right)\right)\left(\partial_0 A_\nu\right) + \left(\partial^0 A^\nu\right)\left(\tfrac{\partial}{\partial(\partial_0 A^\mu)}\left(\partial_0 A_\nu\right)\right)\right] = \\ &= -\tfrac{1}{2}\left[\delta_{\nu\mu}\left(\partial_0 A_\nu\right) + \left(\partial^0 A^\nu\right)\tfrac{\partial}{\partial(\partial_0 A^\mu)}\left(\partial_0 A_\nu\right)\right]. \end{aligned} \tag{W.175}$$

   For the second summand, we use $A_\nu = g_{\nu\kappa}A^\kappa$ and obtain (remember $\partial_0 = \partial^0$)

$$\begin{aligned} \tfrac{\partial}{\partial(\partial_0 A^\mu)}\left(\partial_0 A_\nu\right) &= \tfrac{\partial}{\partial(\partial_0 A^\mu)}\left(\partial_0 g_{\nu\kappa}A^\kappa\right) = \\ &= \tfrac{\partial}{\partial(\partial_0 A^\mu)}\left[\left(\partial_0 g_{\nu\kappa}\right)A^\kappa + g_{\nu\kappa}\partial_0 A^\kappa\right] = g_{\nu\kappa}\tfrac{\partial}{\partial(\partial_0 A^\mu)}\partial_0 A^\kappa = g_{\nu\kappa}\delta_{\mu\kappa} \end{aligned} \tag{W.176}$$

(due to $\partial_0 g_{\nu\kappa} = 0$). Inserting the result gives

$$\begin{aligned} \pi_\mu &= -\tfrac{1}{2}\left[\delta_{\nu\mu}\left(\partial_0 A_\nu\right) + \left(\partial^0 A^\nu\right) g_{\nu\kappa}\delta_{\mu\kappa}\right] = \\ &= -\tfrac{1}{2}\left[\left(\partial_0 A_\mu\right) + \left(\partial^0 A^\nu\right) g_{\nu\mu}\right] = -\tfrac{1}{2}\left[\left(\partial_0 A_\mu\right) + \left(\partial^0 A_\mu\right)\right] = -\left(\partial_0 A_\mu\right). \end{aligned} \tag{W.177}$$

W.7 Operator Ordering

W.7.1 Normal Order

The different Hamilton functions H which we found for the three considered fields (Klein–Gordon, Dirac, radiation) all share the feature of an infinite vacuum energy. We recap the problem on the basis of the Klein–Gordon field. In this case, the Hamilton functions H reads (discrete version)

$$H = \frac{1}{2}\sum_{\mathbf{p}} E_{\mathbf{p}}\left(a^\dagger(\mathbf{p})\, a(\mathbf{p}) + a(\mathbf{p})\, a^\dagger(\mathbf{p})\right). \tag{W.178}$$

By means of the commutation relations $\left[a(\mathbf{p}), a^\dagger(\mathbf{p})\right] = \delta_{\mathbf{pp}'}$, this may be written as

$$H = \sum_{\mathbf{p}} E_{\mathbf{p}}\left(a^\dagger(\mathbf{p})\, a(\mathbf{p}) + \frac{1}{2}\right). \tag{W.179}$$

The second summand, i.e., the sum over the zero point energies $\frac{1}{2}\sum_{\mathbf{p}} E_{\mathbf{p}}$, is *always infinite*. Of course, this is a serious problem. A way out is offered by the fact that physics may take its arguments not only from mathematics alone. One can argue as follows: In the the sum $\sum_{\mathbf{p}}$, there occur arbitrarily large values of $\mathbf{p}$, and thus also arbitrarily large energies. Now it is certainly not sensible (from a physical point of view, not a mathematical one) to take into account energies that may be greater than say the total energy of the universe. In other words, a cut-off at a certain ('large') P is physically legitimate or even necessary, restricting the summation to $|\mathbf{p}| \leq P$. It is not required to specify the exact value of P; it suffices to say that it exists because then the sum over the zero point energies yields a *finite* value which as common reference point for all energies may be neglected. Thus, the world would be in order again.[91]

A formal way to handle the problem of infinities is *normal ordering* (or Wick ordering). The notation[92] is $\mathcal{N}[ab]$ where a and b are scalar field operators. Normal ordering means to rearrange a product of annihilation and creation operators in such

[91] There is another type of infinities in quantum electrodynamics, keyword renormalization, which also is healed by introducing a cut-off, see below.

[92] An alternative notation is enclosing the operator between double-dots, i.e., $:A:$. We prefer here $\mathcal{N}[A]$ due to its better legibility.

a way that

$$\begin{matrix}\text{left hand side} \\ \text{all creation operators}\end{matrix} - \begin{matrix}\text{right hand side} \\ \text{all annihilation operators}\end{matrix} \tag{W.180}$$

thereby *neglecting* all existing commutation relations.

The rearranging depends on whether if we consider bosons (integer spin as with photons or in the Klein–Gordon case) or fermions (half-integer spin as in the Dirac case). For the purpose of a compact notation, we write in this section $\pm$ where the upper sign means bosons and the lower sign fermions. To hold things simple, we use for the commutator $[a, b]$ and the anticommutator $\{a, b\}$ the notation $[a, b]_{\mp}$, i.e., $[a, b]_{-} = ab - ba$ and $[a, b]_{+} = ab + ba$.

Swapping a product of two bosonic operators in normal ordering does not change the sign, but it does so for fermionic operators:

$$\begin{gathered}\text{bosons: } \mathcal{N}\left[aa^{\dagger}\right] = a^{\dagger}a \text{ ; fermions: } \mathcal{N}\left[aa^{\dagger}\right] = -a^{\dagger}a \\ \text{in short } \mathcal{N}\left[aa^{\dagger}\right] = \pm a^{\dagger}a.\end{gathered} \tag{W.181}$$

For example, the operator $a^{\dagger}(\mathbf{p})\, a(\mathbf{p})$ is already in normal order, $\mathcal{N}\left[a^{\dagger}(\mathbf{p})\, a(\mathbf{p})\right] = a^{\dagger}(\mathbf{p})\, a(\mathbf{p})$; the operator $a(\mathbf{p})\, a^{\dagger}(\mathbf{p})$ reads in normal order $\mathcal{N}\left[a(\mathbf{p})\, a^{\dagger}(\mathbf{p})\right] = \pm a^{\dagger}(\mathbf{p})\, a(\mathbf{p})$. If there are several annihilation and creation operators, for instance $ab^{\dagger}c^{\dagger}de$, normal ordering results in $\mathcal{N}\left[ab^{\dagger}c^{\dagger}de\right] = (\pm 1)^{P}\, b^{\dagger}c^{\dagger}ade$ where P gives the number of swappings. We have as an additional rule that the order within the set of destruction and the set of creation operators is unchanged from the original expression. Normal ordering is linear, i.e. (A_i operators, c_i complex numbers): $\mathcal{N}[c_1A_1 + c_2A_2] = c_1\mathcal{N}[A_1] + c_2\mathcal{N}[A_2]$. Note that the normal order of a product differs in general from the product of the normal orders: $\mathcal{N}[A_1A_2] \neq \mathcal{N}[A_1] \cdot \mathcal{N}[A_2]$.

W.7.1.1 Normal Order of Energies and Charges

Energies As an example, we apply the considerations[93] to the three Hamilton functions (see the corresponding sections)

$$\begin{gathered}H_{\text{Klein-Gordon}} = \frac{1}{2}\sum\nolimits_{\mathbf{p}} E_{\mathbf{p}}\left(a^{\dagger}(\mathbf{p})\, a(\mathbf{p}) + a(\mathbf{p})\, a^{\dagger}(\mathbf{p})\right) = \sum\nolimits_{\mathbf{p}} E_{\mathbf{p}}\left(a^{\dagger}(\mathbf{p})\, a(\mathbf{p}) + \frac{1}{2}\right) \\ H_{\text{Dirac}} = \sum\nolimits_{\mathbf{p},r} E_{\mathbf{p}}\left(b_r^{\dagger}(\mathbf{p})\, b_r(\mathbf{p}) - \mathrm{d}_r(\mathbf{p})\, \mathrm{d}_r^{\dagger}(\mathbf{p})\right) = \sum\nolimits_{\mathbf{p},r} E_{\mathbf{p}}\left[b_r^{\dagger}(\mathbf{p})\, b_r(\mathbf{p}) + \mathrm{d}_r^{\dagger}(\mathbf{p})\, \mathrm{d}_r(\mathbf{p}) - 1\right] \\ H_{\text{Radiation}} = \frac{1}{2}\sum\nolimits_{\mathbf{k},r} \omega_{\mathbf{k}}\left(\alpha_r^{\dagger}(\mathbf{k})\, \alpha_r(\mathbf{k}) + \alpha_r(\mathbf{k})\, \alpha_r^{\dagger}(\mathbf{k})\right) = \sum\nolimits_{\mathbf{k},r} \omega_{\mathbf{k}}\left(\alpha_r^{\dagger}(\mathbf{k})\, \alpha_r(\mathbf{k}) + \frac{1}{2}\right).\end{gathered} \tag{W.182}$$

The second form of these Hamiltonians is derived by applying the appropriate (anti-)commutation relations. We see that there is in each case an infinite vacuum

---

[93] We confine ourselves to the discussion of the discrete case; the considerations for the continuous case run analogously.

energy. As stated above, the 'true' energies are given by the normal ordered expressions which implies simply neglecting the infinite vacuum energy term, retaining only the number operators. Thus we have symbolically

$$H_{\text{final}} = \mathcal{N}\left[H_{\text{before}}\right]. \tag{W.183}$$

This yields

$$\begin{aligned} H_{\text{Klein-Gordon}} &= \mathcal{N}\left[\tfrac{1}{2}\sum_{\mathbf{p}} E_{\mathbf{p}}\left(a^{\dagger}(\mathbf{p})\,a(\mathbf{p}) + a(\mathbf{p})\,a^{\dagger}(\mathbf{p})\right)\right] = \sum_{\mathbf{p}} E_{\mathbf{p}} a^{\dagger}(\mathbf{p})\,a(\mathbf{p}) \\ H_{\text{Dirac}} &= \mathcal{N}\left[\sum_{\mathbf{p},r} E_{\mathbf{p}}\left(b_r^{\dagger}(\mathbf{p})\,b_r(\mathbf{p}) - \mathrm{d}_r(\mathbf{p})\,\mathrm{d}_r^{\dagger}(\mathbf{p})\right)\right] = \sum_{\mathbf{p},r} E_{\mathbf{p}}\left(b_r^{\dagger}(\mathbf{p})\,b_r(\mathbf{p}) + \mathrm{d}_r^{\dagger}(\mathbf{p})\,\mathrm{d}_r(\mathbf{p})\right) \\ H_{\text{Radiation}} &= \mathcal{N}\left[\tfrac{1}{2}\sum_{\mathbf{k},r}\omega_{\mathbf{k}}\left(\alpha_r^{\dagger}(\mathbf{k})\,\alpha_r(\mathbf{k}) + \alpha_r(\mathbf{k})\,\alpha_r^{\dagger}(\mathbf{k})\right)\right] = \sum_{\mathbf{k},r}\omega_{\mathbf{k}}\alpha_r^{\dagger}(\mathbf{k})\,\alpha_r(\mathbf{k}). \end{aligned} \tag{W.184}$$

By comparison with (W.182) we see that the annoying infinities have disappeared. It does not matter whether we omit the (infinite) vacuum energy or whether we normal order H. In other words: Normal ordering makes vacuum energy go away.

The above considerations are not restricted to H but hold for all observables. Thus, in general, *any* string of operators in field theory has to be normal ordered to avoid infinities. Without this convention, results are nonsensical, in general.

Note that since *all* observables A are defined a priori as normal ordered products, this fact is often not explicitly mentioned. Thus, there is often no explicit additional notation to mark the normal ordered form. For instance, we have $H = \frac{1}{2}\sum_{\mathbf{p}} E_{\mathbf{p}}\left(a^{\dagger}(\mathbf{p})\,a(\mathbf{p}) + a(\mathbf{p})\,a^{\dagger}(\mathbf{p})\right)$ and its normal ordered form $H = \sum_{\mathbf{p}} E_{\mathbf{p}} a^{\dagger}(\mathbf{p})$ $a(\mathbf{p})$ In many texts, it is $H = \sum_{\mathbf{p}} E_{\mathbf{p}} a^{\dagger}(\mathbf{p})\,a(\mathbf{p})$ which from the start is presented as Hamilton function of the Klein–Gordon field, mostly without further comment and without indicating that it is the normal ordered form.

Charge In addition, we want to consider briefly the total charge Q of the Dirac system. The 4-current density is given by the normal ordered expression[94]

$$j^{\mu}(x) = q\mathcal{N}\left[\bar{\psi}(x)\,\gamma^{\mu}\psi(x)\right] \tag{W.185}$$

where q is the charge of the electron. j^{μ} fulfills the continuity equation

$$\partial_{\mu} j^{\mu}(x) = 0. \tag{W.186}$$

The operator for the total charge $\hat{Q}$ is given by

$$\begin{aligned} \hat{Q} &= \int \mathrm{d}^3x\; j^0(x) = q\int \mathrm{d}^3x\;\mathcal{N}\left[\bar{\psi}(x)\,\gamma^{\mu}\psi(x)\right] = \\ &= q\sum_{\mathbf{p},r}\left(b_r^{\dagger}(\mathbf{p})\,b_r(\mathbf{p}) - \mathrm{d}_r^{\dagger}(\mathbf{p})\,\mathrm{d}_r(\mathbf{p})\right) = q\sum_{\mathbf{p},r}\left(N_{e,r\mathbf{p}} - N_{p,r\mathbf{p}}\right). \end{aligned} \tag{W.187}$$

[94] Remember that the fermionic 4-current (probability current) is given by $\bar{\psi}\gamma^{\mu}\psi$, see Appendix U, Vol. 1.

We know that the 'd-particles', i.e., the positrons, and the 'b -particles', i.e., the electrons, are oppositely charged. This fact is confirmed by considering the eigenvalues of $\hat{Q}$. The commutators of $\hat{Q}$ with the creation operators are given by.

$$\left[\hat{Q}, b_r^\dagger(\mathbf{p})\right] = q b_r^\dagger(\mathbf{p}) \;\; ; \;\; \left[\hat{Q}, \mathrm{d}_r^\dagger(\mathbf{p})\right] = -q \mathrm{d}_r^\dagger(\mathbf{p}) . \tag{W.188}$$

Now assume that we have a state $|\Psi\rangle$ which is eigenstate of $\hat{Q}$ with the eigenvalue Q

$$\hat{Q}|\Psi\rangle = Q|\Psi\rangle . \tag{W.189}$$

It follows with (W.188)

$$\hat{Q}b_r^\dagger(\mathbf{p})|\Psi\rangle = \left(b_r^\dagger(\mathbf{p})\hat{Q} + q b_{r\mathbf{p}}^\dagger\right)|\Psi\rangle = \left(b_r^\dagger(\mathbf{p}) Q + q b_{r\mathbf{p}}^\dagger\right)|\Psi\rangle = (Q+q) b_r^\dagger(\mathbf{p})|\Psi\rangle \tag{W.190}$$

and analogously for the other operators. All in all we have

$$\begin{aligned} \hat{Q}b_r^\dagger(\mathbf{p})|\Psi\rangle &= (Q+q) b_r^\dagger(\mathbf{p})|\Psi\rangle \;\; ; \;\; \hat{Q}\mathrm{d}_r^\dagger(\mathbf{p})|\Psi\rangle = (Q-q)\mathrm{d}_r^\dagger(\mathbf{p})|\Psi\rangle \\ \hat{Q}b_r(\mathbf{p})|\Psi\rangle &= (Q-q_0) b_r(\mathbf{p})|\Psi\rangle \;\; ; \;\; \hat{Q}\mathrm{d}_r(\mathbf{p})|\Psi\rangle = (Q+q)\mathrm{d}_r(\mathbf{p})|\Psi\rangle . \end{aligned} \tag{W.191}$$

Thus, creating an electron or deleting a positron adds q to the total charge; the vacuum has charge zero.

W.7.1.2 Normal Order for General Field Operators

The ground state energy of the normal ordered Hamilton function $H = \sum_{\mathbf{p}} E_{\mathbf{p}} a^\dagger(\mathbf{p}) a(\mathbf{p})$ is zero due to $a(\mathbf{p})|0\rangle = 0$ (or equivalently $\langle 0| a^\dagger(\mathbf{p}) = 0$):

$$\langle 0|\mathcal{N}[H]|0\rangle = \langle 0|\sum_{\mathbf{p}} E_{\mathbf{p}} a^\dagger(\mathbf{p}) a(\mathbf{p})|0\rangle = \langle 0|\sum_{\mathbf{p}} E_{\mathbf{p}} a^\dagger(\mathbf{p})\left(a(\mathbf{p})|0\rangle\right) = 0. \tag{W.192}$$

With creation operators to the left and annihilation operators to the right, *any* normal ordered operator A has a vacuum expectation value of zero. Note that this holds true even if $\langle 0|A|0\rangle \neq 0$, i.e., if the mean value does not vanish:

$$\langle 0|\mathcal{N}[A]|0\rangle = 0 \text{ always.} \tag{W.193}$$

Next we will consider the normal ordering of a product of two general field operators A and B. To this end, we split the field operator A into two parts, one containing all destruction operators A^d, the other containing all creation operators A^c, and $A = A^c + A^d$.[95] Normal ordering the product $A(x)B(y)$ yields

[95]In many textbooks one finds the notation $A = A^+ + A^-$, where the upper index marks the sign of energy, i.e., A^+ contains the destruction operators and A^- the creation operators. But this notation

$$\begin{aligned}\mathcal{N}\left[A(x)B(y)\right] &= \mathcal{N}\left[\left(A^d(x)+A^c(x)\right)\left(B^d(y)+B^c(y)\right)\right] = \\ &= \mathcal{N}\left[A^d(x)B^d(y)\right]+\mathcal{N}\left[A^d(x)B^c(y)\right]+\mathcal{N}\left[A^c(x)B^d(y)\right]+\mathcal{N}\left[A^c(x)B^c(y)\right] = \\ &= A^d(x)B^d(y) \pm B^c(y)A^d(x)+A^c(x)B^d(y)+A^c(x)B^c(y)\,.\end{aligned} \tag{W.194}$$

As is seen, the second term in the last line is the only one with changed order of the operators. So we have immediately for the difference of normal ordered product and product itself

$$\mathcal{N}\left[A(x)B(y)\right] - A(x)B(y) = \pm B^c(y)A^d(x) - A^d(x)B^c(y) = -\left[A^d(x), B^c(y)\right]_{\mp}\,. \tag{W.195}$$

Now the commutator for bosonic operators and the anticommutator for fermionic operators is a c-number,[96] or in other words: the difference $\mathcal{N}\left[A(x)B(y)\right] - A(x)B(y)$ contains no operators anymore. We note that this result is important in the further discussion.

W.7.1.3 Discussion of Normal Ordering

Normal ordering is firmly established in quantum field theory. However, it appears like an arbitrary rule for quantization, a mere ad-hoc convention. The problem is clear to see: the commutation relations are suspended for this step and only for this step. As we have seen, commutation relations are in the heart of Quantum Mechanics, and normal ordering simply overrides these key elements. A theory would be highly desirable which gets along without such an artificial feature.

In classical mechanics, order does not import: px equals xp. In general, products of classical operators can be written in many equivalent ways, and it is not automatically clear which one has to be quantized.[97] However, this is to be expected: the transition from classical mechanics to quantum mechanics *must* necessarily be ambiguous - if it were unambiguous, quantum mechanics would be superfluous.[98] There has to be something new in quantum mechanics. In first quantization, this is the change from Poisson brackets to commutator relations, and in second quantization it is (perhaps) normal ordering.

Whatsoever - the cornerstone of physics is comparison with experimental results. We can accept a physical theory if and only if it agrees with the observations. And

may sometimes be a little bit confusing, especially for beginners, since in A^+ there are the terms $\sim e^{-ikx}$, while the terms $\sim e^{ikx}$ are in A^-. Moreover, a few authors use the signs in the upper index in the reverse meaning. Thus, in order to avoid misunderstanding, we use A^d and A^c instead of A^+ and A^-.

[96] One distinguishes c-numbers (classical numbers) in contrast to q-numbers (quantum mechanical numbers, i.e., operators).

[97] In Vol. 1, we have argued that a classical operator like xp should give a Hermitian quantum operator; so we introduced the Hermitian term $\frac{xp+px}{2}$. But this criterion does not work here, since the operators $aa^\dagger$ and $a^\dagger a$ both are Hermitian.

[98] The transition from classical mechanics to quantum mechanics *cannot* be unambiguous. In contrast, the transition from quantum mechanics to classical mechanics *must* be unambiguous.

the normal ordered Hamilton function results in what is actually observed: In QFT, the *normal ordered* Hamiltonian is the *observable* Hamiltonian.[99] Normal ordering gives a meaningful quantum field theory; if this were not the case, it would not persist. We must accept this as the way nature works. Perhaps a future theory will get rid of this apparent inconsistency.

W.7.2 Time Order

Time ordering[100] $\mathcal{T}$ can be first understood as a practical tool to simplify the notation of complicated series as e.g. the time evolution operator $U_I(t, t_0)$ in the interaction picture. In Appendix Q, Vol. 1, we have found the representation (Dyson series)

$$U_I(t, t_0) = \sum_{n=0}^{\infty} \left(\frac{1}{i\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \ldots \int_{t_0}^{t_{n-1}} dt_n\, H_I(t_1) H_I(t_2) \ldots H_I(t_n)\,. \tag{W.196}$$

Note that the H_I at different times will not commute, $[H_I(t_1), H_I(t_2)] \neq 0$, in general. Thus, the order of time is of great importance; here we have ordered times with $t_0 \leq t_n \leq t_{n-1} \leq \cdots \leq t_2 \leq t_1 \leq t$.

The calculation of this integral is cumbersome, not least because the upper limits of the integrals are all different. To circumvent this difficulty, we introduce an distribution $\Theta(t_1, t_2, \ldots, t_n)$ with the properties

$$\Theta(t_1, t_2, \ldots, t_n) = \begin{cases} 1 & \text{if } t_1 > t_2 > \ldots > t_n \\ 0 & \text{otherwise} \end{cases}\,. \tag{W.197}$$

Then we can write

$$U_I(t, t_0) = \sum_{n=0}^{\infty} \left(\frac{1}{i\hbar}\right)^n \int_{t_0}^{t} dt_1 \ldots \int_{t_0}^{t} dt_n\, \Theta(t_1, \ldots, t_n) H_I(t_1) \ldots H_I(t_n)\,. \tag{W.198}$$

Note that in this version all integrals have as upper limit the *same* value t, since $\Theta(t_1, t_2, \ldots, t_n)$ guarantees that the additional contributions of the integrals vanish - this is the important step. Now for $n = 2$ we have the two possibilities $t_1 > t_2$ and $t_2 > t_1$, for $n = 3$ there are six possibilities ($t_1 > t_2 > t_3$, $t_1 > t_3 > t_2$ and so on), and for arbitrary n we have $n!$ possibilities or permutations. Thus, if we consider all permutations, we can write

---

[99]In other words: a not normal ordered Hamiltonian $\mathcal{H}_{\text{not normal}}$ does not represent an observable. In this sense, $\mathcal{H}_{\text{not normal}}$ is a Hermitian operator, but not an observable. However, in other contexts the zero point energy may be measurable, and that implies a not normal ordered Hamiltonian, as is the case e.g. in molecular vibrations.

[100]Although we need time ordering only in a later section, we consider it here in anticipation on account of its intrinsic proximity to normal ordering.

$$U_I = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} \mathrm{d}t_1 \ldots \int_{t_0}^{t} \mathrm{d}t_n \sum_{\pi \in S_n} \Theta\left(t_{\pi(1)}, \ldots, t_{\pi(n)}\right) H_I\left(t_{\pi(1)}\right) \ldots H_I\left(t_{\pi(n)}\right) \tag{W.199}$$

where $\pi \in S_n$ is one of the $n!$ permutations of the numbers $1, 2, \ldots, n$. Note that (W.199) and (W.196) are strictly identical.

We now define time ordering for a string of operators A_i by

$$\mathcal{T}\left[A_1(t_1), \ldots, A_n(t_n)\right] = \sum_{\pi \in S_n} \Theta\left(t_{\pi(1)}, \ldots, t_{\pi(n)}\right) A_1\left(t_{\pi(1)}\right), \ldots, A_n\left(t_{\pi(n)}\right). \tag{W.200}$$

Due to the properties of Θ as given in (W.197), time ordering picks out exactly that order of operators for which the times are ordered in the right way. Thus, $\mathcal{T}$ guarantees that the operators act in the physical correct order and not the later one before the others.

By means of $\mathcal{T}$, we can now write (W.199) as

$$U_I(t, t_0) = \sum_{n=0}^{\infty} \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} \mathrm{d}t_1 \int_{t_0}^{t} \mathrm{d}t_2 \ldots \int_{t_0}^{t} \mathrm{d}t_n \, \mathcal{T}\left[H_I(t_1) H_I(t_2) \ldots H_I(t_n)\right] \tag{W.201}$$

or more compactly as (*Dyson's series or expansion*)

$$U_I(t, t_0) = \mathcal{T} \exp\left(-\frac{i}{\hbar} \int_{t_0}^{t} \mathrm{d}t' \, H_I\left(t'\right)\right). \tag{W.202}$$

We need this equation and these considerations in the further discussion.

We will need also the time ordered product of two scalar field operators. Let A and B scalar fields, either bosonic (upper sign) or fermionic (lower sign). Then the time ordered product of A and B is defined by

$$\mathcal{T}\left[A(x) B(y)\right] = \begin{cases} A(x) B(y) & \text{for } x^0 > y^0 \\ \pm B(y) A(x) & \phantom{\text{for }} y^0 > x^0 \end{cases}. \tag{W.203}$$

By means of the Heaviside function[101] θ, this may be written more compactly as

$$\mathcal{T}\left[A(x) B(y)\right] = \theta\left(x^0 - y^0\right) A(x) B(y) \pm \theta\left(y^0 - x^0\right) B(y) A(x). \tag{W.205}$$

Similar to the case of normal ordering, we are interested in the difference of a product of two field operators and its time ordered form. It is given by

---

[101] Remember the definition of $\theta(x)$:

$$\theta(x) = \begin{cases} 1 & \text{for } x \geq 0 \\ 0 & \phantom{\text{for }} x < 0 \end{cases} \tag{W.204}$$

$$\mathcal{T}\left[A(x)B(y)\right] - A(x)B(y) = \begin{cases} A(x)B(y) - A(x)B(y) = 0 \\ \pm B(y)A(x) - A(x)B(y) \end{cases} \text{for } \begin{matrix} x^0 > y^0 \\ y^0 > x^0 \end{matrix}. \tag{W.206}$$

With $A(x) = A^d(x) + A^c(x)$, the commutator in the last line is given by

$$\begin{aligned} &\pm B(y)A(x) - A(x)B(y) = -\left[A(x), B(y)\right]_{\mp} = \\ &= -\left[A^d(x) + A^c(x), B^d(y) + B^c(y)\right]_{\mp} = -\left[A^d(x), B^c(y)\right]_{\mp} - \left[A^c(x), B^d(y)\right]_{\mp} \end{aligned} \tag{W.207}$$

due to $\left[A^d(x), B^d(y)\right]_{\mp} = \left[A^c(x), B^c(y)\right]_{\mp} = 0$. It follows finally

$$\mathcal{T}\left[A(x)B(y)\right] - A(x)B(y) = \begin{cases} 0 \\ -\left[A^d(x), B^c(y)\right]_{\mp} - \left[A^c(x), B^d(y)\right]_{\mp} \end{cases} \text{for } \begin{matrix} x^0 > y^0 \\ y^0 > x^0 \end{matrix}. \tag{W.208}$$

As in case of normal ordering, we see that the difference is made of commutators and hence contains no operators, i.e., is a *c*-number.

W.7.3 *Time Ordering and Normal Ordering*

Bringing together time ordering and normal ordering[102] of two scalar fields A and B leads with (W.195) to

$$\mathcal{T}\left[A(x)B(y)\right] - \mathcal{N}\left[A(x)B(y)\right] = T\left[A(x)B(y)\right] - A(x)B(y) + \left[A^d(x), B^c(y)\right]_{\mp}. \tag{W.209}$$

This yields

$$\mathcal{T}\left[A(x)B(y)\right] = \mathcal{N}\left[A(x)B(y)\right] \begin{cases} +\left[A^d(x), B^c(y)\right]_{\mp} \\ -\left[A^c(x), B^d(y)\right]_{\mp} \end{cases} \text{for } \begin{matrix} x^0 > y^0 \\ y^0 > x^0 \end{matrix}. \tag{W.210}$$

As is seen, we can express the time ordered product of two scalar field operators by their normal ordered product plus another term which contains no operators. This result will play an important below.

[102] $\mathcal{T}$ and $\mathcal{N}$ are sometimes called time ordering operator and normal ordering operator. For the sake of good order, we want to point out that this is a misnomer. An operator is an object which, when applied to a state, gives new information, as for instance the angular momentum operator. In this sense, $\mathcal{T}$ and $\mathcal{N}$ are not operators and would be better named instructions. On the other hand, the naming is established and we will use it, too, keeping in mind the caveat.

W.7.4 Exercises and Solutions

1. Prove (W.184).
Solution: Exemplarily, we treat detailed the Dirac case. It is

$$\begin{aligned} H_{\text{Dirac}} = \mathcal{N}\left[\sum\nolimits_{\mathbf{p},r} E_{\mathbf{p}}\left(b_r^\dagger(\mathbf{p})\,b_r(\mathbf{p}) - \mathrm{d}_r(\mathbf{p})\,\mathrm{d}_r^\dagger(\mathbf{p})\right)\right] = \\ = \mathcal{N}\left[\sum\nolimits_{\mathbf{p},r} E_{\mathbf{p}} b_r^\dagger(\mathbf{p})\,b_r(\mathbf{p})\right] - \mathcal{N}\left[\sum\nolimits_{\mathbf{p},r} E\mathrm{d}_r(\mathbf{p})\,\mathrm{d}_r^\dagger(\mathbf{p})\right] = \\ = \sum\nolimits_{\mathbf{p},r} E_{\mathbf{p}}\mathcal{N}\left[b_r^\dagger(\mathbf{p})\,b_r(\mathbf{p})\right] - \sum\nolimits_{\mathbf{p},r} E_{\mathbf{p}}\mathcal{N}\left[\mathrm{d}_r(\mathbf{p})\,\mathrm{d}_r^\dagger(\mathbf{p})\right] = \\ = \sum\nolimits_{\mathbf{p},r} E_{\mathbf{p}} b_r^\dagger(\mathbf{p})\,b_r(\mathbf{p}) + \sum\nolimits_{\mathbf{p},r} E_{\mathbf{p}}\mathrm{d}_r^\dagger(\mathbf{p})\,\mathrm{d}_r(\mathbf{p}) = \\ = \sum\nolimits_{\mathbf{p},r} E_{\mathbf{p}}\left(b_r^\dagger(\mathbf{p})\,b_r(\mathbf{p}) + \mathrm{d}_r^\dagger(\mathbf{p})\,\mathrm{d}_r(\mathbf{p})\right). \end{aligned} \tag{W.211}$$

2. Consider the Klein–Gordon field (discrete version)

$$\phi(x) = \phi^d(x) + \phi^c(x) = \sum_{\mathbf{k}} \frac{1}{\sqrt{2VE_{\mathbf{k}}}}\left(a(\mathbf{k})\,e^{-ikx} + a^\dagger(\mathbf{k})\,e^{ikx}\right). \tag{W.212}$$

Calculate explicitly $\mathcal{N}[\phi(x)\,\phi(y)] - \phi(x)\,\phi(y)$.
Solution: With (W.195) we obtain

$$\mathcal{N}[\phi(x)\,\phi(y)] - \phi(x)\,\phi(y) = -\left[\phi^d(x), \phi^c(y)\right]_- . \tag{W.213}$$

Inserting ϕ^d and ϕ^c yields

$$\begin{aligned} \mathcal{N}[\phi(x)\,\phi(y)] - \phi(x)\,\phi(y) = -\left[\sum\nolimits_{\mathbf{k}} \tfrac{1}{\sqrt{2VE_{\mathbf{k}}}} a(\mathbf{k})\,e^{-ikx}, \sum\nolimits_{\mathbf{k}'} \tfrac{1}{\sqrt{2VE_{\mathbf{k}'}}} a^\dagger(\mathbf{k})\,e^{ik'y}\right]_- = \\ = -\sum\nolimits_{\mathbf{k},\mathbf{k}'} \tfrac{1}{\sqrt{2VE_{\mathbf{k}}}}\tfrac{1}{\sqrt{2VE_{\mathbf{k}'}}} e^{-ikx} e^{ik'y}\left[a(\mathbf{k}), a^\dagger(\mathbf{k})\right]_- = \\ = -\sum\nolimits_{\mathbf{k},\mathbf{k}'} \tfrac{1}{\sqrt{2VE_{\mathbf{k}}}}\tfrac{1}{\sqrt{2VE_{\mathbf{k}'}}} e^{-ikx} e^{ik'y}\delta_{\mathbf{k},\mathbf{k}'} = -\sum\nolimits_{\mathbf{k}} \tfrac{1}{2VE_{\mathbf{k}}} e^{-ik(x-y)}. \end{aligned} \tag{W.214}$$

As is explicitly seen, the result contains no operators.

W.8 Interacting Fields, Quantum Electrodynamics

In order to describe interacting fields (up to now we were considering free fields only), we now combine the pieces which we have developed previously. We describe the interaction by means of the interaction picture which is based on the interaction Hamiltonian $\mathcal{H}_I$. Thereby, we confine our considerations to the interaction of fermions with spin $1/2$ and photons, i.e., to the study of *quantum electrodynamics* (QED). To formulate $\mathcal{H}_I$ we need the Lagrangians $\mathcal{L}$ of the Dirac field and of the radiation field plus a term which couples these two fields. Our focus will be on scattering. The transition amplitude from the incoming to the outgoing state is described by means of the so-called S-matrix which we have, finally, to approximate in a suitable manner to describe scattering processes of lowest orders.

W.8.1 Lagrangian

In order to bring together electrons, positrons and photons, we invoke the principle of minimal coupling. Since we introduced and discussed it already in Appendix T, Vol. 1, we recap it here briefly. The approach replaces the 4-momentum p_μ by the 4-vector $p_\mu - qA_\mu$ (q is the charge of the considered fermion), i.e.,[103]

$$p_\mu \rightarrow p_\mu - qA_\mu \Rightarrow i\partial_\mu \rightarrow i\partial_\mu - qA_\mu. \tag{W.215}$$

It follows for the Lagrangian

$$\mathcal{L}^{\text{Dirac}} = \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi \rightarrow \bar{\psi}\left(\gamma^\mu\left(i\partial_\mu - qA_\mu\right) - m\right)\psi = \bar{\psi}\left(\gamma^\mu i\partial_\mu - m\right)\psi - q\bar{\psi}\gamma^\mu A_\mu\psi. \tag{W.216}$$

This means that by this substitution we have the free Dirac Lagrangian plus an interaction term:

$$\mathcal{L}^{\text{Dirac}} \rightarrow \mathcal{L}^{\text{Dirac}} + \mathcal{L}^{\text{interaction}}\ ;\ \mathcal{L}^{\text{interaction}} = -q\bar{\psi}\gamma^\mu A_\mu\psi. \tag{W.217}$$

The term $q\bar{\psi}\gamma^\mu A_\mu\psi$ is the interface between fermions and photons, containing contributions of both 'worlds'.[104]

In this way, we can write the *total Lagrangian* $\mathcal{L}$ as the sum of the two free Lagrangians (Dirac and photon) and the interaction Lagrangian:

$$\begin{aligned}\mathcal{L} &= \mathcal{L}^{\text{Dirac}} + \mathcal{L}^{\text{photon}} + \mathcal{L}^{\text{interaction}} = \\ &= \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi - \tfrac{1}{2}\left(\partial^\mu A^\nu\right)\left(\partial_\mu A_\nu\right) - q\bar{\psi}\gamma^\mu A_\mu\psi.\end{aligned} \tag{W.218}$$

Thus, $\mathcal{L}$ describes the electromagnetic interaction between electrons, positrons[105] and photons. It is the basic equation for quantum electrodynamics which is one of the best, if not the best, proven theories in physics.

One can just as well regard $q\bar{\psi}\gamma^\mu\psi$ as 4-current j^μ which enters the Lagrangian for electrodynamics in the form $\mathcal{L}^{\text{electrodynamics}} = -\frac{1}{2}\left(\partial^\mu A^\nu\right)\left(\partial_\mu A_\nu\right) + j^\mu A_\mu$ (see section 'Normal ordering').

[103]Remind $p_\mu = i\partial_\mu$, see Appendix T, Vol. 1.

[104]Note that the notation of the charge may cause some confusion since it is not standardized. Here, q means the fermionic charge and the specific charge of the electron ist denoted by $q = -e$. But one finds also the notation e for the general charge and $-e_0$ for the electronic charge. In other contexts, e_0 means a hypothetical, not observable charge of the electron, i.e., the 'bare' charge. So watch out.

[105]Or muons and tauons and their antiparticles.

W.8.2 Conjugated Momentum, Hamiltonian

Next we search for the Hamiltonian density $\mathcal{H}$ of the Lagrangian (W.218). As is seen, in the interaction term $\bar{\psi}\gamma^{\mu}A_{\mu}\psi$ there are no time derivatives $(\partial_0\psi)$ and $(\partial_0 A^{\mu})$ of the fields ψ and A^{μ}. In other words, the conjugated momenta are the same as in the free case, i.e., without interaction, and the Hamiltonian for the interaction reads simply $\mathcal{H}^{\text{interaction}} = -\mathcal{L}^{\text{interaction}}$. Thus, the Hamiltonian is given by

$$\begin{aligned}\mathcal{H} &= \mathcal{H}^{\text{Dirac}} + \mathcal{H}^{\text{photon}} + \mathcal{H}^{\text{interaction}} = \\ &= \bar{\psi}\left(-i\gamma^{\mu}\partial_{\mu} + m\right)\psi - \tfrac{1}{2}\left(\partial^{\mu}A_{\nu}\right)\left(\partial^{\mu}A^{\nu}\right) + q\bar{\psi}\gamma^{\mu}A_{\mu}\psi.\end{aligned} \tag{W.219}$$

We have to add one further step, namely normal ordering. As discussed above in Section 'Operator ordering', operators in quantum field theory have to be normal ordered to be meaningful. Thus, the Hamiltonian for the interaction (which is the term we are interested in for the following) reads in its final version

$$\mathcal{H}^{\text{interaction}} = q\mathcal{N}\left[\bar{\psi}\gamma^{\mu}A_{\mu}\psi\right]. \tag{W.220}$$

Note that ψ and A_{μ} are *free* field operators.

W.8.3 Interaction Picture, Time Evolution Operator

It comes as no surprise that the equation of motions for (W.218) or (W.219) cannot be solved in closed form. Instead, one invokes the interaction picture. Since this issue was introduced and discussed in Appendix Q, Vol. 1, and above in section 'Operator ordering', we recap here only the main points very briefly.

In the interaction picture, it is assumed that the Hamilton function H, as given for instance in the Schrödinger picture, can be written as $H = H_0 + H_1$, where H_0 is the free part and H_1 the interaction part. Usually, H_0 may be solved exactly. Then we define states $|\psi_I(t)\rangle$ and operators $B_I(t)$ in the interaction picture by

$$|\psi_I(t)\rangle = e^{iH_0 t}|\psi(t)\rangle \;\; ; \; B_I(t) = e^{iH_0 t}Be^{-iH_0 t} \tag{W.221}$$

and the time behavior of the state $|\psi_I(t)\rangle$ is given by

$$i\frac{\mathrm{d}}{\mathrm{d}t}|\psi_I(t)\rangle = e^{iH_0 t}H_1 e^{-iH_0 t}|\psi_I(t)\rangle = H_I(t)|\psi_I(t)\rangle\,. \tag{W.222}$$

The time evolution operator $U_I(t, t_0)$ makes contact between $|\psi_I(t_0)\rangle$ and $|\psi_I(t)\rangle$, i.e.,

$$|\psi_I(t)\rangle = U_I(t, t_0)|\psi_I(t_0)\rangle \tag{W.223}$$

and obeys the differential equation

$$i\frac{\mathrm{d}}{\mathrm{d}t}U_I\left(t,t_0\right)=H_I\left(t\right)U_I\left(t,t_0\right) \tag{W.224}$$

The (formal) solution may be written as

$$U_I\left(t,t_0\right)=\mathcal{T}\ \exp\left(-i\int_{t_0}^{t}\mathrm{d}t'H_I\left(t'\right)\right). \tag{W.225}$$

The knowledge of $U_I\left(t,t_0\right)$ enables us to calculate $\left|\psi_I\left(t\right)\right\rangle$ for a given $\left|\psi_I\left(t_0\right)\right\rangle$. Assume that in a certain process the system is at time t_0 in the initial state $\left|\psi_I\left(t_0\right)\right\rangle = \left|i_I\right\rangle$. Then we have $\left|\psi_I\left(t\right)\right\rangle = U_I\left(t,t_0\right)\left|\psi_I\left(t_0\right)\right\rangle$, see (W.223), and the probability P_{fi} to find it at time t in a final state $\left|\psi_I\left(t\right)\right\rangle = \left|f_I\right\rangle$ is given by

$$P_{fi}=\left|\left\langle f_I\right|U_I\left(t,t_0\right)\left|i_I\right\rangle\right|^2. \tag{W.226}$$

Note that transition probabilities are independent from the picture chosen. Denoting the states for $t=t_0$ and t in the Schrödinger picture by $\left|i_S\right\rangle$ and $\left|f_S\right\rangle$, we have shown in Appendix Q, Vol. 1, that $\left\langle f_I\right|U_I\left(t,t_0\right)\left|i_I\right\rangle = \left\langle f_S\right|U_S\left(t,t_0\right)\left|i_S\right\rangle$ - the transition amplitudes in the Schrödinger and the interaction picture are equal.

W.8.4 S-Operator

We now focus our interest upon scattering. In a scattering process, one can distinguish three phases and their idealization:

- Phase 1: At the initial time, the initial particles are widely separated. We idealize this by assuming $t=-\infty$ for the initial time at which we have non-interacting initial particles, i.e., *free* particles.
- Phase 2: The particles encounter each other and interact. Possibly, some of the initial particles are destroyed and new final ones are created. After that, the (new) particles run away from each other.
- Phase 3: At the final time, the final particles are again widely separated. Again, we idealize this by assuming $t=\infty$ for the final time at which we have non interacting final particles, i.e., *free* particles.

The idealization of this process is tantamount to saying that in phases 1 and 3 the interaction is switched off and, in addition, that the 'interaction time' of phase 2 is much shorter than the times needed to run from the source to the scattering center or from the scattering center to the detector. Under these assumptions, we can choose the initial and the final time as $-\infty$ and $+\infty$.

We formalize now this idealized process as seen in the interaction picture. In the beginning we have the initial state $\left|i\right\rangle = \left|\psi_I\left(-\infty\right)\right\rangle$. The time evolution operator $U_I\left(\infty,-\infty\right)$ changes this state into the final state $\left|\psi_I\left(\infty\right)\right\rangle = U_I\left(\infty,-\infty\right)\left|\psi_I\left(-\infty\right)\right\rangle$. Now let be $\left|f\right\rangle$ one of the possible final states. The transition amplitude into this certain final state[106] is given by $\langle f \, |\psi_I(\infty)\rangle = \langle f| \, U_I(\infty, -\infty) \, |i\rangle$; thus, the probability to find this final state $|f\rangle$ for a given initial state $|i\rangle$ is given by $|\langle f \, |\psi_I(\infty)\rangle|^2$.

We now introduce the abbreviating notation

$$S = U_I(\infty, -\infty) \tag{W.227}$$

called *scattering operator* or scattering matrix (or simply S-matrix or S-*operator*; the letter S stems from scattering, of course). Thus, we can write[107]

$$\langle f \, |\psi_I(\infty)\rangle = \langle f| \, U_I(\infty, -\infty) \, |i\rangle = \langle f| \, S \, |i\rangle = S_{fi}. \tag{W.228}$$

With the time evolution operator (W.225) we have

$$S = U_I(\infty, -\infty) = \mathcal{T} \, \exp\left(-i \int_{-\infty}^{\infty} \mathrm{d}t H_I(t)\right) \tag{W.229}$$

and with $H_I(t) = \int \mathrm{d}^4x \; \mathcal{H}_I(x)$

$$S = \mathcal{T}\left[e^{-i\int \mathrm{d}^4x \; \mathcal{H}_I(x)}\right]. \tag{W.230}$$

In the case under consideration, namely quantum electrodynamics, $\mathcal{H}_I$ is given by (W.220) and we have

$$\mathcal{H}_I(x) = \mathcal{H}^{\text{interaction}} = q\mathcal{N}\left[\bar{\psi}\gamma^\mu A_\mu \psi\right]. \tag{W.231}$$

So we have solved, at least in principle, the scattering problem in quantum electrodynamics. Assume an initial state $|i\rangle$. Then the probability to find after the scattering process the final state $|f\rangle$ is given by $|\langle f| \, S \, |i\rangle|^2$ where the S-matrix is given by (W.230) and the interaction Hamiltonian $\mathcal{H}_I(x)$ by (W.231).

W.8.5 Approximating S

Note that (W.230) together with (W.231) is an exact formulation without approximations. We would have finished the problem, if we could find a closed analytical evaluation of the integral $\int \mathrm{d}^4x \; \mathcal{H}_I(x)$ for given $\mathcal{H}_I(x)$. However, such an evaluation does not exist (or pretty much never), and we have to recourse to suitable approximations.

---

[106] Remind that the states $|i\rangle$und $|f\rangle$ are free states (eigenkets of H_0).

[107] Remind that transition amplitudes in the Schrödinger and in the interaction picture are equal.

The usual procedure is to expand the exponential $\mathcal{T}\ \exp\left[-i\int d^4x\ \mathcal{H}_I(x)\right]$ in a series of the form

$$S = \mathcal{T}\left[1 + (-i)\int d^4x\ \mathcal{H}_I(x) + \frac{(-i)^2}{2!}\int\int d^4x\ d^4y\ \mathcal{H}_I(x)\,\mathcal{H}_I(y) + \cdots\right]. \tag{W.232}$$

For many applications, it is sufficient to consider only the first few terms. If $\mathcal{H}_I$ is small enough compared to the full Hamiltonian density $\mathcal{H}$ (which is indeed the case, as we will see[108]), this proceeding will give satisfying results. Thus, the remaining sections are devoted to the discussion of the two terms

$$S^{(1)} = -i\int d^4x\ \left[\mathcal{H}_I(x)\right]\ ;\ S^{(2)} = -\frac{1}{2}\int\int d^4x\ d^4y\ \mathcal{T}\ \left[\mathcal{H}_I(x)\,\mathcal{H}_I(y)\right]. \tag{W.233}$$

The time ordering symbol $\mathcal{T}$ for $S^{(1)}$ may be omitted, since there is only one time to consider.

As we will see, the discussion will be quite extensive, though the two terms look rather simple and innocent. We begin in the next section with $S^{(1)}$. After that, we provide some tools as contractions, propagators and the Wick theorem. They are needed for the discussion of $S^{(2)}$. This discussion is only introductory and exemplary; a thorough and detailed consideration of QED would be beyond the scope of this short introduction.

W.9 *S*-Matrix, First Order

We now want to discuss the first order term $S^{(1)}$

$$S^{(1)} = -i\int d^4x\ \left[\mathcal{H}_I(x)\right] = -iq\int d^4x\ \mathcal{N}\left[\bar{\psi}\not{A}\psi\right] \tag{W.234}$$

in some detail.

With (W.234), we have the simplest case of the scattering matrix, and it is good practice to consider the simplest case first. However, discussing $S^{(1)}$ is not only a convenient finger exercise. It also facilitates the discussion of more complex cases for several reasons. One of them is that $S^{(1)}$ encompasses eight elementary processes which are constituting the more complex cases in higher orders of $S^{(n)}$. A closer look at these elementary processes will provide simple rules which enable us to write down transition amplitudes quite easily.[109] What makes life even easier is the close

[108] In QED, the smallness parameter in $\mathcal{H}_I = q\mathcal{N}\left[\bar{\psi}\not{A}\psi\right]$ is $|q| = |e|$ which in our natural units has the value ∼0.303.

[109] The connection of transition amplitude and scattering cross section will be discussed exemplarily in the context of considering $S^{(2)}$, see below.

connection between those rules and their graphical representation in form of the so-called Feynman diagrams.

To avoid disappointment later on, we want to point out already here that none of the those eight elementary processes can occur for real particles. However, we will learn a lot from them and, as mentioned, they are building blocks of the scattering processes of higher orders.

Note that the processes can not 'really' occur in the way they are described here; but in a different way, they exist for real particles. For one thing, in the case of *external fields* apply different considerations as discussed below, for another thing the processes can exist in higher orders $S^{(n)}$. Take for instance pair annihilation. Below, we consider the 'impossible' first-order case $e^-e^+ \to \gamma$. Of course, 'real' pair annihilation exists, but as $e^-e^+ \to 2\gamma$ which is part of the second order term $S^{(2)}$.

W.9.1 Preliminary Note: Virtual Particles

Real and virtual particles are also called on mass-shell and off mass-shell. Here some comments on their definition.

The inner product of the 4-momentum p is given by $p^2 = p_\mu p^\mu = \left(p^0\right)^2 - \mathbf{p}^2 = E_\mathbf{p}^2 - \mathbf{p}^2 = m^2$. The identity $p^2 = m^2$ is also known as *mass shell condition*. It is the usual dispersion relation for relativistic particles. However, in e.g. scattering processes, there occur particles with $p^2 \neq m^2$, i.e., they exist 'off mass-shell'. These *virtual particles* can not appear in the initial and final states of real processes, but are only emitted and reabsorbed in intermediate steps. One can argue that the energy-time uncertainty relation $\Delta E \Delta t \sim \hbar$ allows for off mass-shell particles with energy ΔE provided they don't live longer than $\Delta t \sim \hbar/\Delta E$. In addition, the finite velocity $\leq c$ of those particles leads to a finite range $\Delta x \lesssim c\hbar/\Delta E$ or in natural units $\Delta x \lesssim 1/\Delta E$. In short: Virtual particles can not be measured; they exist only fleetingly.

Concerning the eight processes we will study now, this means that they cannot be realized in the described form with three on-shell particles.

W.9.2 Field Operators

For clearer distinction, we use from now on the following notation: momentum and spin of fermions[110] are labeled by p and r, momentum and polarization of photons are labeled by k and λ.

[110]The considerations also apply to other fermions as muons and tauons.

The field operators are given in the sections 'Quantization of free fields' for the Dirac and the photon field. The continuous versions read

$$\begin{gathered}\psi(x)=\sum_r\int \mathrm{d}^3p\,\sqrt{\frac{m}{(2\pi)^3E_{\mathbf{p}}}}\left(b_r(\mathbf{p})\,u_r(\mathbf{p})\,e^{-ipx}+\mathrm{d}_r^{\dagger}(\mathbf{p})\,w_r(\mathbf{p})\,e^{ipx}\right)\\ A^{\mu}(x)=\int\sum_\lambda \mathrm{d}^3k\,\sqrt{\frac{1}{(2\pi)^3E_{\mathbf{k}}}}\epsilon_\lambda^{\mu}(\mathbf{k})\left(\alpha_\lambda(\mathbf{k})\,e^{-ikx}+\alpha_\lambda^{\dagger}(\mathbf{k})\,e^{ikx}\right).\end{gathered} \tag{W.235}$$

The commutation relations are given by

$$\begin{gathered}\left\{b_r(\mathbf{p}),b_{r'}^{\dagger}(\mathbf{p}')\right\}=\delta_{rr'}\delta\left(\mathbf{p}-\mathbf{p}'\right)\;;\;\left\{\mathrm{d}_r(\mathbf{p}),\mathrm{d}_{r'}^{\dagger}(\mathbf{p}')\right\}=\delta_{rr'}\delta\left(\mathbf{p}-\mathbf{p}'\right)\\ \left[\alpha_\lambda(\mathbf{k}),\alpha_{\lambda'}^{\dagger}(\mathbf{k}')\right]=\delta_{\lambda\lambda'}\delta^3\left(\mathbf{k}-\mathbf{k}'\right).\end{gathered} \tag{W.236}$$

All other (anti-)commutators vanish.

In view of normal ordering, it is advantageous for some considerations to aggregate the contributions of deletion and creation operators in the form

$$\begin{gathered}\psi(x)=\psi^d(x)+\psi^c(x)\;;\;\bar\psi(x)=\bar\psi^d(x)+\bar\psi^c(x)\\ A^{\mu}(x)=A^{\mu d}(x)+A^{\mu c}(x).\end{gathered} \tag{W.237}$$

Thereby, the creation and annihilation parts of the field operators are explicitly given by[111]

$$\begin{gathered}\psi^d(x)=\sum_r\int \mathrm{d}^3p\,\sqrt{\frac{m}{(2\pi)^3E_{\mathbf{p}}}}b_r(\mathbf{p})\,u_r(\mathbf{p})\,e^{-ipx}\;;\;\psi^c(x)=\sum_r\int \mathrm{d}^3p\,\sqrt{\frac{m}{(2\pi)^3E_{\mathbf{p}}}}\mathrm{d}_r^{\dagger}(\mathbf{p})\,w_r(\mathbf{p})\,e^{ipx}\\ \bar\psi^d(x)=\sum_r\int \mathrm{d}^3p\,\sqrt{\frac{m}{(2\pi)^3E_{\mathbf{p}}}}\mathrm{d}_r(\mathbf{p})\,\bar w_r(\mathbf{p})\,e^{-ipx}\;;\;\bar\psi^c(x)=\sum_r\int \mathrm{d}^3p\,\sqrt{\frac{m}{(2\pi)^3E_{\mathbf{p}}}}b_r^{\dagger}(\mathbf{p})\,\bar u_r(\mathbf{p})\,e^{ipx}\\ A^{\mu d}(x)=\sum_\lambda\int \mathrm{d}^3k\,\sqrt{\frac{1}{(2\pi)^3E_{\mathbf{k}}}}\epsilon_\lambda^{\mu}(\mathbf{k})\,\alpha_\lambda(\mathbf{k})\,e^{-ikx}\;;\;A^{\mu c}(x)=\sum_\lambda\int \mathrm{d}^3k\,\sqrt{\frac{1}{(2\pi)^3E_{\mathbf{k}}}}\epsilon_\lambda^{\mu}(\mathbf{k})\,\alpha_\lambda^{\dagger}(\mathbf{k})\,e^{ikx}.\end{gathered} \tag{W.238}$$

$\psi^d(x)$ contains all terms $\sim e^{-ipx}$ and all annihilation operators, $\psi^c(x)$ all terms $\sim e^{ipx}$ and all creation operators; analogously for $A^{\mu}(x)$.

The action of these operators is given by

$$\left.\begin{matrix}\psi^d\\ \bar\psi^d\\ \not{A}^d\end{matrix}\right\}\text{ annihilates a(n) }\left\{\begin{matrix}\text{electron}\\ \text{positron}\\ \text{photon}\end{matrix}\right.\;;\;\left.\begin{matrix}\psi^c\\ \bar\psi^c\\ \not{A}^c\end{matrix}\right\}\text{ creates a(n) }\left\{\begin{matrix}\text{positron}\\ \text{electron}\\ \text{photon}\end{matrix}\right.. \tag{W.239}$$

Thus, $\psi=\psi^d+\psi^c$ annihilates an electron and creates a positron, whereas $\bar\psi=\bar\psi^d+\bar\psi^c$ annihilates a positron and creates an electron.

[111] As mentioned in the section 'Operator ordering', many textbooks write ψ^+ for the annihilation part and ψ^- for the creation part. We use ψ^d and ψ^c instead of ψ^+ and ψ^-, where d stands for 'destroying' and c für 'creating'.

Table W.3 The eight possible processes of $\mathcal{H}_I$

Term	Normal order	Description
$\bar{\psi}^d \not{A}^d \psi^d$		$\gamma e^- e^+ \to$ vacuum
$\bar{\psi}^d \not{A}^d \psi^c$	$\mathcal{N}\left[\bar{\psi}^d_\alpha \gamma^\mu_{\alpha\beta} A^+_\mu \psi^c_\beta\right] = -\psi^c_\beta \bar{\psi}^d_\alpha \gamma^\mu_{\alpha\beta} A^d_\mu$	$\gamma e^+ \to e^+$
$\bar{\psi}^d \not{A}^c \psi^d$	$A^c_\mu \psi^d \gamma^\mu \psi^d$	$e^- e^+ \to \gamma$
$\bar{\psi}^d \not{A}^c \psi^c$	$\mathcal{N}\left[\bar{\psi}^d_\alpha \gamma^\mu_{\alpha\beta} A^c_\mu \psi^c_\beta\right] = -\gamma^\mu_{\alpha\beta} A^c_\mu \psi^c_\beta \bar{\psi}^d_\alpha$	$e^+ \to \gamma e^+$
$\bar{\psi}^c \not{A}^d \psi^d$		$\gamma e^- \to e^-$
$\bar{\psi}^c \not{A}^d \psi^c$		$\gamma \to e^- e^+$
$\bar{\psi}^c \not{A}^c \psi^d$		$e^- \to \gamma e^-$
$\bar{\psi}^c \not{A}^c \psi^c$		vacuum $\to \gamma e^- e^+$

W.9.3 *Eight Elementary Processes of $\mathcal{H}_I$*

Let us look which processes are allowed by the interaction Hamiltonian

$$\mathcal{H}_I(x) = q\mathcal{N}\left[\bar{\psi}(x)\not{A}(x)\psi(x)\right]. \tag{W.240}$$

We insert (W.237) into this equation, expand the brackets and obtain the following $2^3 = 8$ terms

$$\mathcal{H}_I(x) = q\mathcal{N}\begin{bmatrix} \bar{\psi}^d\not{A}^d\psi^d + \bar{\psi}^d\not{A}^d\psi^c + \psi^d\not{A}^c\psi^d + \bar{\psi}^d\not{A}^c\psi^c + \\ +\bar{\psi}^c\not{A}^d\psi^d + \bar{\psi}^c\not{A}^d\psi^c + \bar{\psi}^c\not{A}^c\psi^d + \bar{\psi}^c\not{A}^c\psi^c \end{bmatrix}. \tag{W.241}$$

Consider for instance the term $\bar{\psi}^c\not{A}^d\psi^c$ which already is in normal order. Following the action as given in (W.239), it creates a positron (ψ^c), annihilates a photon ($\not{A}^d$) and creates an electron ($\bar{\psi}^c$), or in short $\gamma \to e^- e^+$ (pair production). Before discussing the action of the other seven terms we note that only three terms are not in normal order (which means upper index c to the left, d to the right), namely $\bar{\psi}^d\not{A}^d\psi^c$, $\bar{\psi}^d\not{A}^c\psi^d$ and $\bar{\psi}^d\not{A}^c\psi^c$:

$$\mathcal{H}_I(x) = q\begin{bmatrix} \bar{\psi}^d\not{A}^d\psi^d + \mathcal{N}\left[\bar{\psi}^d\not{A}^d\psi^c\right] + \mathcal{N}\left[\bar{\psi}^d\not{A}^c\psi^d\right] + \mathcal{N}\left[\bar{\psi}^d\not{A}^c\psi^c\right] + \\ +\bar{\psi}^c\not{A}^d\psi^d + \bar{\psi}^c\not{A}^d\psi^c + \bar{\psi}^c\not{A}^c\psi^d + \bar{\psi}^c\not{A}^c\psi^c \end{bmatrix}. \tag{W.242}$$

Normal ordering for instance the first of these terms, $\bar{\psi}^d\not{A}^d\psi^c$, can be achieved as follows:

$$\mathcal{N}\left[\bar{\psi}^d\not{A}^d\psi^c\right] = \mathcal{N}\left[\bar{\psi}^d\gamma^\mu A^d_\mu\psi^c\right] = \mathcal{N}\left[\bar{\psi}^d_\alpha\gamma^\mu_{\alpha\beta}A^d_\mu\psi^c_\beta\right] = -\psi^c_\beta\bar{\psi}^d_\alpha\gamma^\mu_{\alpha\beta}A^d_\mu \tag{W.243}$$

where α and β denote the entries of the 4-spinors ψ^d and ψ^c; $\gamma^\mu_{\alpha\beta}$ is the element $(\alpha\beta)$ of the matrix γ^μ. We sum up all processes in Table W.3.

Note that one of the premises of quantum field theory is fulfilled: we have different numbers and types of particles in the incoming and outgoing channel.

One can show that none of these eight processes can be realized with on-shell particles only. The discussion is much easier when using Feynman diagrams, see below and the exercises.

W.9.4 Two Worked Out Examples

Exemplarily, we want to consider in the following two of the eight processes in more detail, namely 1) $\bar{\psi}^c A\!\!\!/^c \psi^d$ or $e^- \rightarrow \gamma e^-$ (emission of a photon) and 2) $\bar{\psi}^c A\!\!\!/^d \psi^c$ or $\gamma \rightarrow e^- e^+$ (pair production). Thereby, we make use of the field operators in the continuous version as given in (W.235).

W.9.4.1 First Example: Emission of a Photon, $e^- \rightarrow \gamma e^-$

The S-matrix element reads

$$S^{(1)} = -iq \int \mathrm{d}^4 x \; \bar{\psi}^c A\!\!\!/^c \psi^d. \tag{W.244}$$

Initial and final states The initial and final states are given by

$$|i\rangle = b_R^\dagger (\mathbf{P}) |0\rangle \;\; ; \;\; |f\rangle = b_{R'}^\dagger \left(\mathbf{P}'\right) \alpha_\Lambda^\dagger (\mathbf{K}) |0\rangle \rightarrow \langle f| = \langle 0| \alpha_\Lambda (\mathbf{K})\, b_{R'} \left(\mathbf{P}'\right) \tag{W.245}$$

i.e., an incoming electron with quantum numbers R and $\mathbf{P}$ and an outgoing electron with $\left(R', \mathbf{P}'\right)$ plus an outgoing photon with $(\Lambda, \mathbf{K})$.

Matching rules Let us first consider $\psi^d |i\rangle$. We have with continuous field operators[112]

$$\begin{aligned} \psi^d |i\rangle &= \sum_r \int \mathrm{d}^3 p \sqrt{\tfrac{m}{(2\pi)^3 E_\mathbf{p}}} b_r (\mathbf{p})\, u_r (\mathbf{p})\, e^{-ipx} b_R^\dagger (\mathbf{P}) |0\rangle = \\ &= \sum_r \int \mathrm{d}^3 p \sqrt{\tfrac{m}{(2\pi)^3 E_\mathbf{p}}} u_r (\mathbf{p})\, e^{-ipx} \delta_{rR} \delta (\mathbf{p} - \mathbf{P}) |0\rangle = \sqrt{\tfrac{m}{(2\pi)^3 E_\mathbf{P}}} u_R (\mathbf{P})\, e^{-iPx} |0\rangle \end{aligned} \tag{W.246}$$

or in short

$$\psi^d b_R^\dagger (\mathbf{P}) |0\rangle = \sqrt{\frac{m}{(2\pi)^3 E_\mathbf{P}}} u_R (\mathbf{P})\, e^{-iPx} |0\rangle . \tag{W.247}$$

The argument runs as follows: $b_R^\dagger (\mathbf{P}) |0\rangle$ creates an electron with quantum numbers $\mathbf{P}$ and R. The only annihilation operator $b_r (\mathbf{p})$ which can destroy this electron

[112]For the discrete version, replace $\sqrt{\frac{m}{(2\pi)^3 E_\mathbf{p}}} \rightarrow \sqrt{\frac{m}{V E_\mathbf{p}}}$ and $\sqrt{\frac{1}{2(2\pi)^3 E_\mathbf{K}}} \rightarrow \sqrt{\frac{1}{2 V E_\mathbf{K}}}$.

Table W.4 Incoming and outgoing contributions

$\psi^d b_R^\dagger(\mathbf{P})\|0\rangle = \sqrt{\frac{m}{(2\pi)^3 E_\mathbf{P}}} u_R(\mathbf{P})\, e^{-iPx}\|0\rangle$	$A^{\mu d}(x)\,\alpha_\Lambda^\dagger(\mathbf{K})\|0\rangle = \sqrt{\frac{1}{2(2\pi)^3 E_\mathbf{K}}}\epsilon_\Lambda^\mu(\mathbf{K})\, e^{-iKx}\|0\rangle$
$\langle 0\| \bar\psi^c b_R(\mathbf{P}) = \langle 0\| \sqrt{\frac{m}{(2\pi)^3 E_\mathbf{P}}}\bar u_R(\mathbf{P})\, e^{iPx}$	$\langle 0\| \alpha_\Lambda(\mathbf{K}) A^{\mu c}(x) = \langle 0\| \sqrt{\frac{1}{2(2\pi)^3 E_\mathbf{K}}}\epsilon_\Lambda^\mu(\mathbf{K})\, e^{iKx}$

has to have the same quantum numbers; if not, it acts onto the vacuum, yielding zero. If one prefers a more formal argument, one considers $b_r(\mathbf{p})\, b_R^\dagger(\mathbf{P})$. Due to the anticommutation rule $\left\{b_r(\mathbf{p}), b_{r'}^\dagger(\mathbf{p}')\right\} = \delta_{rr'}\delta\left(\mathbf{p}-\mathbf{p}'\right)$ we have

$$\left\{b_r(\mathbf{p}), b_R^\dagger(\mathbf{P})\right\} = \delta_{rR}\delta(\mathbf{p}-\mathbf{P}) \rightarrow b_r(\mathbf{p})\, b_R^\dagger(\mathbf{P}) = \delta_{rR}\delta(\mathbf{p}-\mathbf{P}) - b_R^\dagger(\mathbf{P})\, b_r(\mathbf{p})\,. \tag{W.248}$$

Thus, for $b_r(\mathbf{p})\, b_R^\dagger(\mathbf{P})|0\rangle$ follows

$$b_r(\mathbf{p})\, b_R^\dagger(\mathbf{P})|0\rangle = \delta_{rR}\delta(\mathbf{p}-\mathbf{P})|0\rangle - b_R^\dagger(\mathbf{P})\, b_r(\mathbf{p})|0\rangle = \delta_{rR}\delta(\mathbf{p}-\mathbf{P})|0\rangle \tag{W.249}$$

due to $b_r(\mathbf{p})|0\rangle = 0$.

In other words, only those parts of the field operator ψ^d survive which match the quantum numbers of the incoming particle. The same holds for outgoing particles: for $\bar\psi^c$ only the term $b_{R'}^\dagger\left(\mathbf{P}'\right)$ contributes, and for $A\!\!\!/^c$ only the term $\alpha_\Lambda(\mathbf{K})$. Written as a short formula or rule we have in general

Transition amplitude In this way we arrive at

$$\begin{gathered}\langle f|\, S^{(1)}\,|i\rangle = -iq \int \mathrm{d}^4x\ \langle f|\, \bar\psi^c A\!\!\!/^c \psi^d\,|i\rangle = \\ = -iq \int \mathrm{d}^4x \sqrt{\frac{m}{(2\pi)^3 E_{\mathbf{P}'}}}\bar u_{R'}\left(\mathbf{P}'\right) e^{iP'x}\sqrt{\frac{1}{2(2\pi)^3 E_\mathbf{k}}}\gamma_\mu\epsilon_\Lambda^\mu(\mathbf{K})\, e^{iKx}\sqrt{\frac{m}{(2\pi)^3 E_\mathbf{P}}} u_R(\mathbf{P})\, e^{-iPx} = \\ = -iq\sqrt{\frac{m}{(2\pi)^3 E_{\mathbf{P}'}}}\bar u_{R'}\left(\mathbf{P}'\right)\sqrt{\frac{1}{2(2\pi)^3 E_\mathbf{k}}}\gamma_\mu\epsilon_\Lambda^\mu(\mathbf{K})\sqrt{\frac{m}{(2\pi)^3 E_\mathbf{P}}} u_R(\mathbf{P}) \int \mathrm{d}^4x\ e^{iP'x} e^{iKx} e^{-iPx}.\end{gathered} \tag{W.250}$$

The x-integration[113] yields the four-dimensional delta function $(2\pi)^4\,\delta^{(4)}\left(P'+K-P\right)$ and it follows

$$\begin{gathered}\langle f|\, S^{(1)}\,|i\rangle = -iq\,(2\pi)^4\,\delta^{(4)}\left(P'+K-P\right)\cdot \\ \cdot\sqrt{\frac{m}{(2\pi)^3 E_{\mathbf{P}'}}\frac{1}{2(2\pi)^3 E_\mathbf{k}}\frac{m}{(2\pi)^3 E_\mathbf{P}}}\ \bar u_{R'}\left(\mathbf{P}'\right)\gamma_\mu\epsilon_\Lambda^\mu(\mathbf{K})\, u_R(\mathbf{P})\,.\end{gathered} \tag{W.251}$$

As mentioned above, all processes of first order can not occur with on-shell particles. In the case under consideration, the argument runs as follows: the four-dimensional delta function contains the conservation of the momentum and of the energy. Conservation of the momentum means $\mathbf{P} = \mathbf{P}' + \mathbf{K}$ or $\mathbf{P}' = \mathbf{P} - \mathbf{K}$. Then conservation of the energy reads

[113] Remember $\int \mathrm{d}^4x\ e^{ipx} = (2\pi)^4\,\delta^{(4)}(p)$.

$$E_{\mathbf{P}-\mathbf{K}} \stackrel{!}{=} E_{\mathbf{P}} - E_{\mathbf{K}} \to \sqrt{(\mathbf{P}-\mathbf{K})^2 + m^2} \stackrel{!}{=} \sqrt{\mathbf{P}^2 + m^2} - \sqrt{\mathbf{K}^2}. \tag{W.252}$$

Squaring both sides yields

$$\begin{aligned} (\mathbf{P}-\mathbf{K})^2 + m^2 &\stackrel{!}{=} \mathbf{P}^2 + m^2 - 2\sqrt{\mathbf{K}^2}\sqrt{\mathbf{P}^2 + m^2} + \mathbf{K}^2 \\ &\to \mathbf{PK} \stackrel{!}{=} \sqrt{\mathbf{K}^2}\sqrt{\mathbf{P}^2 + m^2}. \end{aligned} \tag{W.253}$$

But this equation can not be fulfilled since $\mathbf{PK} \leq |\mathbf{P}|\,|\mathbf{K}|$ and $\sqrt{\mathbf{K}^2}\sqrt{\mathbf{P}^2 + m^2} = |\mathbf{K}|\,\sqrt{|\mathbf{P}|^2 + m^2} > |\mathbf{K}|\,|\mathbf{P}|$.

So we conclude that the process as described here is indeed not possible for real fermions and photons, i.e., for on-shell particles. But it can exist e.g. in the frame of a higher order $S^{(n)}$ as we will see in discussion of processes of $S^{(2)}$.

W.9.4.2 Second Example: Pair Production, $\gamma \to e^- e^+$

The S-matrix element reads

$$S^{(1)} = -iq \int \mathrm{d}^4x \; \bar{\psi}^c \not{A}^d \psi^c = -iq \int \mathrm{d}^4x \; \bar{\psi}^c \gamma_\mu \psi^c A^{\mu d}. \tag{W.254}$$

This means that the initial state is a photon; the final state consists of an electron and a positron. Thus, the initial and final states are given by

$$|i\rangle = \alpha_\Lambda^\dagger\,(\mathbf{K})\,|0\rangle \;\; ; \;\; |f\rangle = b_{R_1}^\dagger\,(\mathbf{P}_1)\,\mathrm{d}_{R_2}^\dagger\,(\mathbf{P}_2)\,|0\rangle \to \langle f| = \langle 0|\,b_{R_1}\,(\mathbf{P}_1)\,\mathrm{d}_{R_2}\,(\mathbf{P}_2)\,. \tag{W.255}$$

Following the ‘matching rules’, developed above, the contributions to $\langle f|\,S^{(1)}\,|i\rangle$ for this process are $\sqrt{\frac{1}{(2\pi)^3 E_{\mathbf{k}}}}\epsilon_\Lambda^\mu\,(\mathbf{K})\,e^{-iKx}$ by the photon, whereas the outgoing electron brings $\sqrt{\frac{m}{(2\pi)^3 E_{\mathbf{P}_1}}}\bar{u}_{R_1}\,(\mathbf{P}_1)\,e^{iP_1x}$ and the outgoing positron contributes $\sqrt{\frac{m}{(2\pi)^3 E_{\mathbf{P}_2}}}w_{R_2}\,(\mathbf{P}_2)\,e^{iP_2x}$. It follows

$$\begin{aligned} \langle f|\,S^{(1)}\,|i\rangle &= -iq\,(2\pi)^4\,\delta^{(4)}\,(P_1 + P_2 - K)\,\cdot \\ &\cdot \sqrt{\frac{m}{(2\pi)^3 E_{\mathbf{P}_1}}\frac{1}{2(2\pi)^3 E_{\mathbf{k}}}\frac{m}{(2\pi)^3 E_{\mathbf{P}_2}}}\;\bar{u}_{R_1}\,(\mathbf{P}_1)\,\gamma_\mu \epsilon_\Lambda^\mu\,(\mathbf{K})\,w_{R_2}\,(\mathbf{P}_2)\,. \end{aligned} \tag{W.256}$$

Here the delta function yields:

$$E_{\mathbf{k}} = E_{\mathbf{P}_1} + E_{\mathbf{P}_2} \;\; ; \;\; \mathbf{K} = \mathbf{P}_1 + \mathbf{P}_2. \tag{W.257}$$

Hence

$$|\mathbf{K}| = \sqrt{\mathbf{P}_1^2 + m^2} + \sqrt{\mathbf{P}_2^2 + m^2}\;; \;\; \mathbf{K} = \mathbf{P}_1 + \mathbf{P}_2. \tag{W.258}$$

But this is a contradiction: from the first equation follows $|\mathbf{K}|^2 > (|\mathbf{P}_1| + |\mathbf{P}_2|)^2 + 2m^2$, and from the second $|\mathbf{K}|^2 \leq (|\mathbf{P}_1| + |\mathbf{P}_2|)^2$. Thus, pair production as described here cannot occur. But this is to be expected - we know, that electron-positron pair production with one photon in free space can only occur near e.g. a nucleus which receives some recoil.

W.9.5 External Fields

As mentioned in the introduction to this section, the eight processes cannot be realized with on-shell particles. However, with external fields, the situation is different. We will very briefly sketch the reason. Assume we have a static external field $A^{\mu}_{\text{ext}}(\mathbf{x})$. For the sake of simplicity, we confine the considerations to a scalar potential $V(\mathbf{x})$ which can be e.g. the Coulomb potential of a nucleus.

Scattering of an electron in this external field can be described by replacing the photonic contribution $\not{A}$ by the external field, i.e., by

$$\langle f|\, S^{(1)}\,|i\rangle = -iq\int \mathrm{d}^4x\ \langle f|\,\bar{\psi}^c \not{A}^c \psi^d\,|i\rangle \to \langle f|\, S^{(1)}_{\text{ext}}\,|i\rangle = -iq\int \mathrm{d}^4x\ \langle f|\,\bar{\psi}^c \not{A}^c_{\text{ext}} \psi^d\,|i\rangle\,. \tag{W.259}$$

Assume that we have an incoming electron $(\mathbf{P}, R)$ which is scattered to $(\mathbf{P}', R')$. We compare this with the transition amplitude (W.250) for $e^- \to \gamma e^-$. We have

$$\left\{\begin{matrix} \langle f|\,S^{(1)}\,|i\rangle \\ \langle f|\,S^{(1)}_{\text{ext}}\,|i\rangle \end{matrix}\right\} = -iq\int \mathrm{d}^4x\,\sqrt{\frac{m}{(2\pi)^3\,E_{\mathbf{P}'}}}\bar{u}_{R'}\left(\mathbf{P}'\right)e^{iP'x}\left\{\begin{matrix} \sqrt{\frac{1}{2(2\pi)^3E_{\mathbf{k}}}}\gamma_\mu\epsilon^{\mu}_{\Lambda}\left(\mathbf{K}\right)e^{iKx} \\ \gamma_0 V\left(\mathbf{x}\right) \end{matrix}\right\}\sqrt{\frac{m}{(2\pi)^3\,E_{\mathbf{P}}}}u_R\left(\mathbf{P}\right)e^{-iPx}. \tag{W.260}$$

The term $\gamma_0 V(\mathbf{x})$ has no time dependence. Hence, performing the x-integral gives

$$\int \mathrm{d}^4x\ e^{iP'x}V(\mathbf{x})e^{-iPx} = 2\pi\delta^{(1)}\left(p_0 - p_0'\right)\cdot\tilde{V}\left(\mathbf{p} - \mathbf{p}'\right) \tag{W.261}$$

where $\tilde{V}$ is the Fourier transform of $V(x)$. Now the trouble with virtual particles in $\langle f|\,S^{(1)}\,|i\rangle$ comes from the delta function $\delta^{(4)}\left(P' + K - P\right)$ in (W.251). But for the external field we have instead an one-dimensional delta function $\delta^{(1)}\left(p_0 - p_0'\right)$ which does not pose any difficulties. This is the crucial point.

Due to lack of space, we will not go into greater detail (which is one the many gaps of this text mentioned in the introduction).

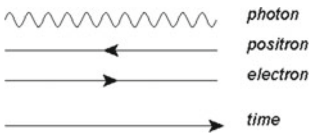


Fig. W.1 Basic lines

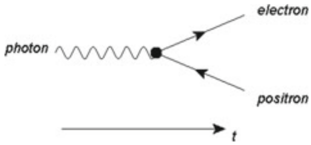


Fig. W.2 Pair production

W.9.6 Feynman Diagrams

By means of *Feynman diagrams*, scattering processes can be visualized in a very nice way. What is more, they also allow for formulating precisely the corresponding expressions for the transition amplitudes.

The diagrams consist of points (called vertices) and lines affixed to the vertices. Fermions are represented by a solid line with an arrow, and photons by a wavy line, see Fig. W.1. External fields are marked by a small cross as in Fig. W.4. Time is placed on the vertical or the horizontal axis; both versions are common. For particles as electrons or muons, the arrow is oriented in direction of the time; for antiparticles as positrons in opposite direction, corresponding to the conception of antiparticles in the Stückelberg-Feynman interpretation.

A further ingredient of the diagrams are vertices, i.e., points where fermions and photons interact.[114] In QED, there are always three lines attached to a vertex: one photonic line and two fermionic lines, one with the arrow toward the vertex and the other with the arrow away from the vertex. As an example, we show in Fig. W.2 the diagram for pair production. Note that in this process an electron and a positron are created (reverse direction of the positron).

The inclination of the lines has no physical meaning, as outlined in Fig. W.3. What matters is the relation to the time axis. A remark on notation: In many cases, when the situation is obvious, there is no labeling - a wavy line is a photon, and electron and positron are clearly identified by their arrows with respect to the time axis. Labeling makes sense e.g. when there are different types of particles and antiparticles, as electrons and muons.

Note that all possible processes of $S^{(1)}$ have *one* vertex. Thus, in the language of Feynman diagrams one can state ‘a single vertex is not physical’.[115]

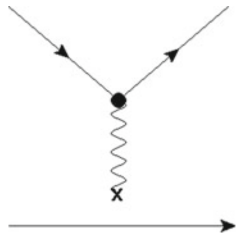
As wee will see below in more detail, Feynman diagrams are precise graphical representations of amplitudes for particle reactions. However, they do not claim to

[114]Note that the order of the series expansion of S gives the number of vertices: diagrams for $S^{(1)}$ have one vertex, for $S^{(2)}$ two vertices and so on.

[115]With external fields as in Fig. W.4, the situation is different. However, in a certain sense, the cross is comparable to a vertex.

Fig. W.3 Inclination of the lines does not matter. Pair production and Compton scattering (see section 'S-matrix, 2. order')

Fig. W.4 Electron scattering in an external field



be precise descriptions images of reality - indeed, they are rather schematic and reduced to the essentials. They show point-like objects which interact at points. But we know that there are waves which interact over a region. For example, in some Feynman diagrams, fermionic or photonic lines are drawn in such a way that they suggest instantaneous propagation (or at least faster than light). But this is an artefact, due to the simplified and schematic representation of a more complex reality. In this respect, the pictures are to be understood with caution - they represent the reality only symbolically.

But they are excellent tools concerning the mathematical formulation of the scattering processes. They greatly reduce the computations; accordingly, they are, in a certain sense, the most common method of computing amplitudes in quantum field theory.

W.9.7 First Feynman Rules

Even for the two simple processes we have just considered we had to do some algebra, and similar calculations in higher order of approximations, i.e., $S^{(2)}$, $S^{(3)}$, …, require an significantly higher amount of algebra. Thus, one is interested in shortcuts similar to the 'matching rules' we found in (W.4), e.g. that an incoming electron with quantum numbers[116] $(R, \mathbf{P})$ produces a factor $\sqrt{\frac{m}{(2\pi)^3 E_P}} u_R(\mathbf{P})\, e^{-iPx}$.

To get more information we consider how this rules are reflected in the transition amplitudes for emission of a photon (W.251) and for pair production (W.256). We write them one above the other:

[116]Remember that the indication of $\mathbf{P}$ fixes the 4-vector P due to $P_0 = \sqrt{\mathbf{P}^2 + m^2}$.

Table W.5 Formal and graphical assignments to particles and vertex; Feynman rules (to be completed)

incoming			outgoing		
e^-	$u_r(\mathbf{p})$		e^-	$\bar{u}_r(\mathbf{p})$	
e^+	$\bar{w}_r(\mathbf{p})$		e^+	$w_r(\mathbf{p})$	
γ	$\epsilon^\mu_\lambda(\mathbf{k})$		γ	$\epsilon^\mu_\lambda(\mathbf{k})$	
		vertex		$-iq\gamma_\mu$	

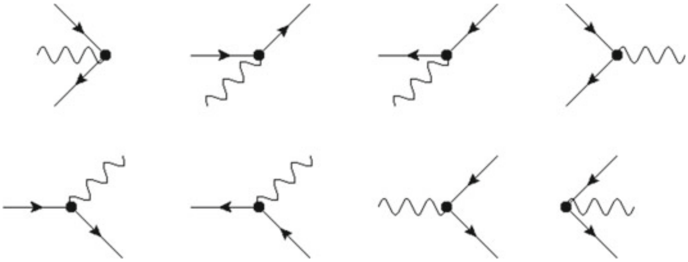


Fig. W.5 The eight processes of $S^{(1)}$. Top: $\gamma e^- e^+ \to$ vacuum, $\gamma e^- \to e^-$, $\gamma e^+ \to e^+$, $e^- e^+ \to \gamma$. Bottom: $e^- \to \gamma e^-$, $e^+ \to \gamma e^+$, $\gamma \to e^- e^+$, vacuum $\to \gamma e^- e^+$

$$\langle f|\, S^{(1)}\, |i\rangle_{\text{emission}} = -iq\,(2\pi)^4\,\delta^{(4)}\left(P' + K - P\right) \cdot \sqrt{\frac{m}{(2\pi)^3 E_{\mathbf{P}'}}\frac{1}{2(2\pi)^3 E_{\mathbf{k}}}\frac{m}{(2\pi)^3 E_{\mathbf{P}}}}\,\bar{u}_{R'}\left(\mathbf{P}'\right)\gamma_\mu\epsilon^\mu_\Lambda(\mathbf{K})\,u_R(\mathbf{P})$$
$$\langle f|\, S^{(1)}\, |i\rangle_{\text{pair}} = -iq\,(2\pi)^4\,\delta^{(4)}\left(P_1 + P_2 - K\right) \cdot \sqrt{\frac{m}{(2\pi)^3 E_{\mathbf{P}_1}}\frac{1}{2(2\pi)^3 E_{\mathbf{k}}}\frac{m}{(2\pi)^3 E_{\mathbf{P}_2}}}\,\bar{u}_{R_1}(\mathbf{P}_1)\,\gamma_\mu\epsilon^\mu_\Lambda(\mathbf{K})\,w_{R_2}(\mathbf{P}_2)\,. \tag{W.262}$$

As is seen, the structure is very similar. From right to left we have the incoming fermion, characterized by spin vectors like $u_R(\mathbf{P})$, then the photon with polarization vector $\epsilon^\mu_\Lambda(\mathbf{K})$, and finally the outgoing fermion. It follows a square root as prefactor which contains the masses and energies of the three items. And finally we have a term which we can write as $(2\pi)^4\,\delta^{(4)}\,(P_{\text{in}} - P_{\text{out}}) \cdot \left(-iq\gamma_\mu\right)$ where P_{in} and P_{out} are the total momenta in the incoming and the outgoing channel. In this way, we can formulate all transition amplitudes of the eight elementary processes of Table W.3 without getting involved into long calculations doing sums and integrals.

What makes life even easier, is the connection with the Feynman diagrams. Indeed, there is a one-to-one correspondence between transition amplitude and Feynman diagram. Each term in (W.262) has a direct counterpart in the diagram. We compile the results in the following table. These assignments are part of what is called *Feynman rules* which (as enhanced version of out matching rules) are the one-to-one translation rules between the Feynman amplitudes in their mathematical form and the representation of the processes in the Feynman diagrams.

And finally, we show in Fig. W.5 the Feynman diagrams for the eight elementary processes.

W.9.8 Exercises and Solutions

1. Show that all eight processes can not occur with real particles.
   Solution: We focus on the conservation of the 4-momentum. As is seen from Fig. W.5, the top and bottom processes are time-inverted versions of each other. If a process can not occur, the same holds for its time-reflection. So we need to consider only one row, say bottom. The last process (vacuum $\to \gamma e^- e^+$) is forbidden due to energy conservation, $E_i = 0, E_f > 0$. The first two processes in the bottom row, $e^- \to \gamma e^-$ and $e^+ \to \gamma e^+$, are equal with regard to conservation of the 4-momentum. Thus, we need to consider only one of them. This means, that there are left $e^- \to \gamma e^-$ and $\gamma \to e^- e^+$. But we have shown above that these processes cannot occur with real particles, see (W.252) and (W.257). Thus, this holds for all processes in the bottom row, and, hence, in the top row.
2. Formulate $\langle f | S^{(1)} | i \rangle$ for all eight processes of $S^{(1)}$ (cf. Fig. W.5).
   Solution: We consider $\gamma e^- \to e^-$. Incoming electron $u_R(\mathbf{P})$, photon $\epsilon^\mu_\Lambda(\mathbf{K})$, outgoing electron $u_{R'}(\mathbf{P}')$. It follows

$$\begin{gathered}\langle f | S^{(1)} | i \rangle = (2\pi)^4 \delta^{(4)}\left(P + K - P'\right) \cdot \\ \sqrt{\frac{m}{(2\pi)^3 E_{\mathbf{P}'}} \frac{1}{2(2\pi)^3 E_{\mathbf{k}}} \frac{m}{(2\pi)^3 E_{\mathbf{P}}}}\, \bar{u}_{R'}\left(\mathbf{P}'\right)\left(-iq\gamma_\mu\right)\epsilon^\mu_\Lambda(\mathbf{K})\, u_R(\mathbf{P}) .\end{gathered} \tag{W.263}$$

   The other cases analogously.

W.10 Contraction, Propagator, Wick's Theorem

The discussion of the second order term $S^{(2)} = -\frac{1}{2}\int\int d^4x\, d^4y\, T\ [\mathcal{H}_I(x)\,\mathcal{H}_I(y)]$ with $\mathcal{H}_I(x) = q\mathcal{N}\left[\bar{\psi}\not{A}\psi\right]$ is quite elaborate and complex compared to $S^{(1)}$. Essentially, this is due to two facts: 1) there are more terms now than in $S^{(1)}$; 2) we have to time-order a product of normal ordered strings of operators. The issue requires new tools which we will provide in this section.

We start with the contraction, i.e., the difference of a time ordered and a normal ordered product of two operators. It will turn out that this term equals essentially the propagator which in turn is closely related to the evolution of a system. Finally, Wick's theorem describes how to get rid of time ordered strings of field operators by means of contractions and (easy to handle) normal ordered products.

W.10.1 Contraction

Given two field scalar operators $A(x)$ and $B(y)$. Then the *contraction* of A and B, written as $\overbrace{A(x)B(y)}$, is defined by the difference of the time ordered and normal ordered product of AB:

$$\contraction{}{A}{(x)}{B}A(x)B(y) = \mathcal{T}\left[A\left(x\right)B\left(y\right)\right] - \mathcal{N}\left[A\left(x\right)B\left(y\right)\right]. \tag{W.264}$$

We have determined this expression in section 'Operator ordering' (see (W.210)); it is given by[117]

$$\contraction{}{A}{(x)}{B}A(x)B(y) = \begin{cases} \left[A^d\left(x\right), B^c\left(y\right)\right]_{\mp} \\ -\left[A^c\left(x\right), B^d\left(y\right)\right]_{\mp} \end{cases} \text{for } \begin{matrix} x^0 > y^0 \\ y^0 > x^0 \end{matrix}. \tag{W.265}$$

Note that performing the (anti)commutators 'uses up' the creation and deletion operators. Thus, the contraction $\contraction{}{A}{(x)}{B}A(x)B(y)$ contains no creation and deletion operators and is a c-number in this regard; we have with $\langle 0\,|0\rangle = 1$

$$\langle 0|\,\contraction{}{A}{(x)}{B}A(x)B(y)\,|0\rangle = \contraction{}{A}{(x)}{B}A(x)B(y). \tag{W.266}$$

Equation (W.265) may be written as[118]

$$\contraction{}{A}{(x)}{B}A(x)B(y) = \theta\left(x^0 - y^0\right)\left[A^d\left(x\right), B^c\left(y\right)\right]_{\mp} - \theta\left(y^0 - x^0\right)\left[A^c\left(x\right), B^d\left(y\right)\right]_{\mp} \tag{W.268}$$

or in more detail:

$$\begin{aligned} \text{bosons: } \contraction{}{A}{(x)}{B}A(x)B(y) &= \theta\left(x^0 - y^0\right)\left[A^d\left(x\right), B^c\left(y\right)\right] + \theta\left(y^0 - x^0\right)\left[B^d\left(y\right), A^c\left(x\right)\right] \\ \text{fermions: } \contraction{}{A}{(x)}{B}A(x)B(y) &= \theta\left(x^0 - y^0\right)\left\{A^d\left(x\right), B^c\left(y\right)\right\} - \theta\left(y^0 - x^0\right)\left\{B^d\left(y\right), A^c\left(x\right)\right\}. \end{aligned} \tag{W.269}$$

We now want to form the vacuum expectation value of the contraction (W.265), i.e., $\langle 0|\,\mathcal{T}\left[A\left(x\right)B\left(y\right)\right] - \mathcal{N}\left[A\left(x\right)B\left(y\right)\right]|0\rangle = \langle 0|\,\contraction{}{A}{(x)}{B}A(x)B(y)\,|0\rangle = A(x)B(y)$. Since the vacuum expectation value of a normal ordered expression always vanishes, we get with (W.266)

$$\langle 0|\,\mathcal{T}\left[A\left(x\right)B\left(y\right)\right]|0\rangle = \contraction{}{A}{(x)}{B}A(x)B(y) \tag{W.270}$$

and with (W.265) follows

[117] Remember $[A, B]_- = AB - BA$ (bosons) and $[A, B]_+ = \{A, B\} = AB + BA$ (fermions).

[118] Remember the definition of the Heaviside function $\theta\left(x\right)$:

$$\theta\left(x\right) = \begin{cases} 1 \\ 0 \end{cases} \text{for } \begin{matrix} x \geq 0 \\ x < 0 \end{matrix} \tag{W.267}$$

$$\langle 0|\,\mathcal{T}\left[A\left(x\right)B\left(y\right)\right]|0\rangle = \begin{cases} \left[A^{d}\left(x\right), B^{c}\left(y\right)\right]_{\mp} & \text{for } x^{0} > y^{0} \\ -\left[A^{c}\left(x\right), B^{d}\left(y\right)\right]_{\mp} & \phantom{\text{for }} y^{0} > x^{0} \end{cases}. \tag{W.271}$$

Remember that the commutators are c-numbers which do not contain any creation or annihilation operators. As we will see in the following section, they play an important role in the further discussion of QED.

Let us draw up an interim balance. We started with the definition of the contraction as $\overbracket{A(x)B}(y) = \mathcal{T}[AB] - \mathcal{N}[AB]$. Conversely, we can express the time ordered product by the sum of normal ordered product plus contraction, $\mathcal{T}[AB] = \mathcal{N}[AB] + \overbracket{A(x)B}(y)$. Of course, this switch only makes sense, if we can find a suitable alternative expression for the contraction. Then we would have indeed achieved our aim to replace the cumbersome time ordering by the the easy to handle normal ordering plus terms which are also rather simple to determine. We will see now that these terms are propagators.

W.10.2 Propagators

W.10.2.1 Green's Function, Propagator

For the sake of simplicity and brevity, we confine ourselves in the following to the Klein–Gordon field. For the Dirac field and the photon field, the results are reported at the end of this section.

Assume we have a system S in some state $|\Sigma\rangle$ which we want to probe with an extra particle p. In an idealized process, we bring S and p into contact at a spacetime point $\left(y^0, \mathbf{y}\right)$ and there will be some interaction between S and p. After that, we will remove p at a spacetime point $\left(x^0, \mathbf{x}\right)$ with $x^0 > y^0$. One may ask if S has remained in its previous state $|\Sigma\rangle$. The answer is given by the projection of (p created at y) $|\Sigma\rangle$ onto (p annihilated at x) $|\Sigma\rangle$, i.e., by the probability amplitude $\langle\Sigma|$ (p annihilated at x) (p created at y) $|\Sigma\rangle$. This amplitude is called Green's function or *propagator* and noted by $G^+(x, y)$. With the notation $\phi^\dagger$ and ϕ for the creation and annihilation operator for p we can write

$$G^{+}\left(x, y\right) = \theta\left(x^{0} - y^{0}\right)\ \langle\Sigma|\,\phi\left(x\right)\phi^{\dagger}\left(y\right)|\Sigma\rangle\,. \tag{W.272}$$

The index $+$ on the left denotes that p is created at y *before* it is destroyed at x; the Heaviside function θ guarantees this behavior. G^+ is also called retarded propagator.

We now invoke the interpretation of Feynman-Stückelberg of antiparticles (see Appendix U, Vol. 1) which sees essentially antiparticles as particles traveling backwards in time. Thus, for the antiparticle $\bar{p}$ we have the corresponding process of creating it at x and then annihilating it at y with $y^0 > x^0$. Since $\phi^\dagger$ and ϕ are the annihilation and creation operators for $\bar{p}$, we can describe this process by

$$G^{-}(x,y) = \theta\left(y^0 - x^0\right)\ \langle\Sigma|\,\phi^{\dagger}(y)\,\phi(x)\,|\Sigma\rangle\,. \tag{W.273}$$

G^{-} is also called advanced propagator.

In QFT, we have to consider particles and antiparticles on an equal footing. It was Feynman who has pointed out that as a consequence we have to consider *both* Green's functions, i.e.,

$$\begin{gathered} G_F(x,y) = G^{+}(x,y) + G^{-}(x,y) = \\ = \theta\left(x^0 - y^0\right)\ \langle\Sigma|\,\phi(x)\,\phi^{\dagger}(y)\,|\Sigma\rangle + \theta\left(y^0 - x^0\right)\ \langle\Sigma|\,\phi^{\dagger}(y)\,\phi(x)\,|\Sigma\rangle \\ \rightarrow G_F(x,y) = \langle\Sigma|\,\mathcal{T}\left[\phi(x)\,\phi^{\dagger}(y)\right]|\Sigma\rangle\,. \end{gathered} \tag{W.274}$$

Note the occurrence of the time ordering operator $\mathcal{T}$. As we will see, such expressions are of particular importance in QFT, and so is time ordering.

Now assume that the state $|\Sigma\rangle$ is given by the vacuum state $|0\rangle$. Then by $\phi^{\dagger}(y)\,|0\rangle$ we create a particle at y and nothing else happens (there is no further interaction). By $\langle 0|\,\phi(x)$, this particle is destroyed at x. Thus, the propagator $\langle 0|\,\phi(x)\,\phi^{\dagger}(y)\,|0\rangle$ describes a particle which propagates from y to x, and $\langle 0|\,\mathcal{T}\left[\phi(x)\,\phi^{\dagger}(y)\right]|0\rangle$ describes this situation for the particle *and* the antiparticle. It is called *(free) Feynman propagator*[119] $\Delta_F(x,y)$:

$$\begin{gathered} i\,\Delta_F(x,y) = \langle 0|\,\mathcal{T}\left[\phi(x)\,\phi^{\dagger}(y)\right]|0\rangle = \\ = \theta\left(x^0 - y^0\right)\ \langle 0|\,\phi(x)\,\phi^{\dagger}(y)\,|0\rangle + \theta\left(y^0 - x^0\right)\ \langle 0|\,\phi^{\dagger}(y)\,\phi(x)\,|0\rangle\,. \end{gathered} \tag{W.275}$$

Let us point out once more that $\theta\left(x^0 - y^0\right)\ \langle 0|\,\phi(x)\,\phi^{\dagger}(y)\,|0\rangle$ describes a particle travelling from y to x for $x^0 > y^0$, and $\theta\left(y^0 - x^0\right)\ \langle 0|\,\phi^{\dagger}(y)\,\phi(x)\,|0\rangle$ describes an antiparticle travelling from x to y for $y^0 > x^0$. Thus, the Feynman propagator is the transition amplitude for two processes, namely (1) a particle p is created at y and annihilated at x, (2) an antiparticle $\bar{p}$ is created at x and annihilated at y. *Both* processes are included in the Feynman propagator; it contains both amplitudes corresponding to the possibility that we have a particle or an antiparticle.

W.10.2.2 Connection with the Contraction

Between the Feynman propagator (W.275) and the contraction (W.265) exists a very close connection - they are identical up to a factor i:

$$i\,\Delta_F(x,y) = \langle 0|\,\overbracket{\phi(x)\,\phi^{\dagger}}(y)\,|0\rangle\,. \tag{W.276}$$

Indeed, from (W.275) we have[120]

---

[119]The factor i is inserted in the definition because it makes things easier in the following. It is missing in some textbooks.

[120]Remind that the vacuum expectation value of a normal ordered product vanishes, $\langle 0|\,\mathcal{N}\left[\phi(x)\,\phi^{\dagger}(y)\right]|0\rangle = 0$.

$$\begin{aligned} i\Delta_F\left(x, y\right) &= \langle 0|\,\mathcal{T}\left[\phi\left(x\right)\phi^{\dagger}\left(y\right)\right]|0\rangle = \\ &= \langle 0|\,\mathcal{T}\left[\phi\left(x\right)\phi^{\dagger}\left(y\right)\right] - \mathcal{N}\left[\phi\left(x\right)\phi^{\dagger}\left(y\right)\right]|0\rangle = \langle 0|\,\overbracket{\phi\left(x\right)\phi^{\dagger}}\left(y\right)|0\rangle\,. \end{aligned} \tag{W.277}$$

In terms of bosonic deletion and creation operators, i.e., $\phi\left(x\right) = \phi^{d}\left(x\right) + \phi^{c}\left(x\right)$, $\phi^{\dagger}\left(y\right) = \phi^{\dagger d}\left(y\right) + \phi^{\dagger c}\left(y\right)$ we have

$$\langle 0|\,\overbracket{\phi\left(x\right)\phi^{\dagger}}\left(y\right)|0\rangle = \theta\left(x^0 - y^0\right)\langle 0|\left[\phi^{d}\left(x\right), \phi^{\dagger c}\left(y\right)\right]|0\rangle + \theta\left(y^0 - x^0\right)\langle 0|\left[\phi^{\dagger d}\left(y\right), \phi^{c}\left(x\right)\right]|0\rangle\,. \tag{W.278}$$

W.10.2.3 Explicit Calculation of the Propagator $\Delta_F\left(x, y\right)$

How looks $\Delta_F\left(x, y\right)$ explicitly? The answer reads[121]

$$\begin{aligned} i\Delta_F\left(x, y\right) &= \langle 0|\,\mathcal{T}\left[\phi\left(x\right)\phi^{\dagger}\left(y\right)\right]|0\rangle = \\ &= \frac{1}{(2\pi)^3}\int\frac{\mathrm{d}^3 p}{\sqrt{2E_{\mathbf{p}}}}\left[\theta\left(x^0 - y^0\right)\,e^{-ip(x-y)} + \theta\left(y^0 - x^0\right)\,e^{ip(x-y)}\right]. \end{aligned} \tag{W.280}$$

Another usual formulation of the free propagator requires complex analysis which is beyond the scope of this short introduction into QFT. Thus, we report here only the result[122]; it reads

$$i\Delta_F\left(x, y\right) = \langle 0|\,\mathcal{T}\left[\phi\left(x\right)\phi^{\dagger}\left(y\right)\right]|0\rangle = \frac{1}{(2\pi)^4}\int \mathrm{d}^4 p\,\frac{i}{p^2 - m^2 + i\varepsilon}e^{-ip(x-y)} \tag{W.282}$$

where $\varepsilon > 0$ is infinitesimal. With the Fourier transform

$$\Delta_F\left(x, y\right) = \frac{1}{(2\pi)^4}\int \mathrm{d}^4 p\,\tilde{\Delta}_F\left(p\right)e^{-ip(x-y)} \tag{W.283}$$

we arrive at the Fourier representation $\tilde{\Delta}\left(p\right)$ of the Feynman propagator for a particle with momentum p (or briefly the momentum space propagator):

$$\tilde{\Delta}_F\left(p\right) = \frac{1}{p^2 - m^2 + i\varepsilon}. \tag{W.284}$$

---

[121] Note that we have

$$\left(\partial^2 + m^2\right)\Delta\left(x, y\right) = -i\delta^4\left(x - y\right) \tag{W.279}$$

Thus, $\Delta\left(x, y\right)$ is evidently Green's function.

[122] One inserts the definition of the Heaviside function

$$\theta\left(x^0 - y^0\right) = \frac{i}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}z\,\frac{e^{-iz\left(x^0 - y^0\right)}}{z + i\varepsilon} \tag{W.281}$$

into (W.280) and obtains (W.282) after some manipulations.

One sees that there is a singularity for $p^2 = E_{\mathbf{p}}^2 - \mathbf{p}^2 = m^2$, i.e., if the usual dispersion relation for a relativistic particle is fulfilled,[123] or, in other words, if the particle is on-shell. But the Feynman propagator encompasses off-shell particles, i.e., virtual particles, which do not obey the relativistic dispersion relation. Such particles are allowed, but only for a short time, as is explained in the previous section.

W.10.2.4 Propagators for the Klein–Gordon Field, the Dirac Field, the Radiation Field

For the Klein–Gordon field $\phi(x)$, the Dirac field $\psi(x)$ and the radiation field $A^\mu(x)$, we have the following contractions:

$$\begin{aligned} \overbracket{\phi(x)\,\phi^\dagger}(y) &= i\Delta_F(x-y) \\ \overbracket{\psi_\alpha(x)\,\bar{\psi}_\beta}(y) = -\overbracket{\bar{\psi}_\beta(y)\,\psi_\alpha}(x) &= iS_{F\alpha\beta}(x-y) \\ \overbracket{A^\mu(x)\,A^\nu}(y) &= i\mathrm{d}_F^{\mu\nu}(x-y)\,. \end{aligned} \tag{W.285}$$

All other contractions vanish, for instance $\overbracket{\psi(x)\,A^\mu}(y) = 0$ and so on. In other words: contractions vanish except when they equal a Feynman propagator.

The propagators are explicitly given by the following expressions:

$$\begin{aligned} \Delta_F(x-y) &= \tfrac{1}{(2\pi)^4}\int \mathrm{d}^4p\, \tfrac{1}{p^2-m^2+i\varepsilon} e^{-ip(x-y)} \\ S_F(x-y) &= \tfrac{1}{(2\pi)^4}\int \mathrm{d}^4p\, \tfrac{\not{p}+m}{p^2-m^2+i\varepsilon} e^{-ip(x-y)} \\ D_F^{\mu\nu}(x-y) &= -g^{\mu\nu}\tfrac{1}{(2\pi)^4}\int \mathrm{d}^4k\, \tfrac{1}{k^2+i\varepsilon} e^{-ik(x-y)}. \end{aligned} \tag{W.286}$$

Three remarks: (1) The Dirac case is also written as

$$S_{F\alpha\beta}(x-y) = \frac{1}{(2\pi)^4}\int \mathrm{d}^4p\, \frac{(\not{p}+m)_{\alpha\beta}}{p^2-m^2+i\varepsilon} e^{-ip(x-y)} \tag{W.287}$$

where $(\not{p}+m)_{\alpha\beta}$ is the $\alpha\beta$-entry of the matrix $\not{p}+m$. Note that the propagators S_F and D_F are matrix-valued. (2) In some textbooks, the infinitesimal character of ε is stressed by writing e.g.

$$\Delta_F(x-y) = \lim_{\varepsilon\to 0^+} \frac{1}{(2\pi)^4}\int \mathrm{d}^4p\, \frac{1}{p^2-m^2+i\varepsilon} e^{-ip(x-y)}. \tag{W.288}$$

(3) In some textbooks, the fraction $\frac{\not{p}+m}{p^2-m^2+i\varepsilon}$ is formally reduced to $\frac{1}{\not{p}-m+i\varepsilon}$. Of course, this is just a symbolical notation; division by a matrix is not defined in this form.

[123] The infinitesimal term $i\varepsilon$ guarantees that one never hits this singularity (pole) exactly; it has to do with the integration procedures in complex analysis.

Fig. W.6 Electron propagator $S_F(p)$ and photon propagator $D_F^{\mu\nu}(k)$ in Feynman diagrams

Finally, in momentum space the propagators read

$$\begin{aligned} \Delta_F(p) &= \tfrac{1}{p^2-m^2+i\varepsilon} \\ S_F(p) &= \tfrac{\not{p}+m}{p^2-m^2+i\varepsilon} \\ D_F^{\mu\nu}(k) &= -g^{\mu\nu}\tfrac{1}{k^2+i\varepsilon}. \end{aligned} \tag{W.289}$$

In Feynman diagrams, propagators are 'clamped' between vertices, see Fig. W.6.

W.10.3 Wick's Theorem

We have introduced two ordering prescriptions - normal ordering $\mathcal{N}$ and time ordering $\mathcal{T}$. Though the justification for normal ordering of operators seems perhaps opaque and not really satisfying, the mathematical procedure itself is simple and offers no difficulties. On the other hand, time ordering is not a bit opaque on the formal level, but the mathematical procedure itself is difficult and very cumbersome.

It would be a big relief and help, if we could express the demanding time ordering of a string of operators by means of the easy to handle normal ordering. In view of the fact, that these two ordering procedures are physically totally different, it comes perhaps as a surprise - but it works, as shown by *Wick's theorem*[124]: we can express time ordering of a string by a combination of normal ordered strings and contractions.

The basic idea for the theorem is time ordering of two scalar field operators A and B equals the sum of the normal ordered product plus the contraction, see (W.264):

$$\mathcal{T}[AB] = \mathcal{N}[AB] + \overset{\sqcap}{AB}. \tag{W.290}$$

The question is what happens if we have to time-order a string of more than two field operators, i.e., $\mathcal{T}[ABCD\ldots]$. The answer is provided by Wick's theorem which we state here without proof.[125]

Wick's theorem: The time ordered product of field operators equals the sum of the normal ordered products, whereby the operators are connected by all possible contractions.[126]

---

[124] Wick, Gian Carlo, 1909–1992; Italian physicist.

[125] The theorem may be proved by induction.

[126] Remember the linearity of normal ordering, $\mathcal{N}[(A+B)C] = \mathcal{N}[AC] + \mathcal{N}[BC]$, and $\mathcal{N}[cA] = c\mathcal{N}[A]$, where c is an expression which does not contain annihilation or creation operators.

$$\mathcal{T}\left[ABC\ldots Z\right] = \mathcal{N}\left[ABC\ldots Z + \text{all possible contractions of } ABC\ldots Z\right]. \tag{W.291}$$

As an example, consider $\mathcal{T}[ABCD]$. We have

$$\begin{aligned} \mathcal{T}[ABCD] &= \mathcal{N}[ABCD] + \\ &+\mathcal{N}\left[\wick{\c A \c B C D}\right] + \mathcal{N}\left[\wick{\c A B \c C D}\right] + \mathcal{N}\left[\wick{\c A B C \c D}\right] + \\ &+\mathcal{N}\left[\wick{A \c B \c C D}\right] + \mathcal{N}\left[\wick{A \c B C \c D}\right] + \mathcal{N}\left[\wick{A B \c C \c D}\right] + \\ &+\wick{\c A \c B \c C \c D} + \wick{\c1 A \c2 B \c1 C \c2 D}. \end{aligned} \tag{W.292}$$

So we can present the single term $\mathcal{T}[ABCD]$ by nine terms, one without, six with one and two with two contractions. This increase in terms may seem annoying, but it actually facilitates the discussion substantially, since normal ordering is a easy to perform procedure in contrast to time ordering. In addition, contractions contain no field operators and hence are c-numbers in this context, i.e., can be moved out of scalar products.

As an illustration, we consider the vacuum expectation value $\langle 0|\mathcal{T}[ABCD]|0\rangle$. Remember that for normal ordered strings always holds $\langle 0|\mathcal{N}[\ldots]|0\rangle = 0$. This means, that all products with a normal ordered part vanish, e.g.,

$$\langle 0|\mathcal{N}\left[\wick{\c A \c B C D}\right]|0\rangle = \wick{\c A \c B}\cdot\langle 0|\mathcal{N}[CD]|0\rangle = 0. \tag{W.293}$$

Thus, only the pure contraction terms survive, and due to $\langle 0|\wick{\c A \c B \c C \c D}|0\rangle = \wick{\c A \c B \c C \c D}\cdot\langle 0\,|0\rangle = \wick{\c A \c B \c C \c D}$ we get the final result

$$\langle 0|\mathcal{T}[ABCD]|0\rangle = \wick{\c A \c B \c C \c D} + \wick{\c1 A \c2 B \c1 C \c2 D} \tag{W.294}$$

where on the r.h.s there are only known terms, essentially propagators.

Wick's theorem provides also an answer to the question how to treat terms like $\mathcal{T}[\mathcal{N}[AB]\,\mathcal{N}[CD]]$ (called mixed time ordered products). We state the answer without proof: A mixed time ordered product equals the sum of the normal ordered products whereby only contractions occur which connect operators from different normal ordered products.

As an example we consider $\mathcal{T}[A_1\mathcal{N}[A_2A_3A_4]]$. We have

$$\begin{aligned} &\mathcal{T}[A_1\cdot\mathcal{N}[A_2A_3A_4]] = \\ &= \mathcal{N}[A_1A_2A_3A_4] + \mathcal{N}\left[\wick{\c A_1 \c A_2 A_3 A_4}\right] + \mathcal{N}\left[\wick{\c A_1 A_2 \c A_3 A_4}\right] + \mathcal{N}\left[\wick{\c A_1 A_2 A_3 \c A_4}\right]. \end{aligned} \tag{W.295}$$

We see immediately, that the vacuum expectation value $\langle 0| \mathcal{T} [A_1 \cdot \mathcal{N} [A_2 A_3 A_4]] |0\rangle$ vanishes since there are only normal ordered terms on the r.h.s.

By Wick's theorem, the difficult time ordering is replaced by the simpler normal ordering and contractions, but there may emerge a remarkable number of those terms. Consider an expression which in some sense is quite close to QED, namely $\mathcal{T} [\mathcal{H} (x_1) \mathcal{H} (x_2)]$ with $\mathcal{H} (x_1) = \mathcal{N} [A (x_1) B (x_1) C (x_1)] = \mathcal{N} [A_1 B_1 C_1]$. We have

$$\begin{aligned} &\mathcal{T} [\mathcal{N} [A_1 B_1 C_1] \cdot \mathcal{N} [A_2 B_2 C_2]] = \mathcal{N} [(A_1 B_1 C_1) (A_2 B_2 C_2)] + \\ &+\mathcal{N} \left[\wick{(\c1 A_1 B_1 C_1)(\c1 A_2 B_2 C_2)} \right] + \text{ 8 other terms with one contraction } + \\ &+\mathcal{N} \left[\wick{(\c1 A_1 \c2 B_1 C_1)(\c1 A_2 \c2 B_2 C_2)} \right] + \text{ 35 other terms with two contractions } + \\ &+\wick{(\c1 A_1 \c2 B_1 \c3 C_1)(\c1 A_2 \c2 B_2 \c3 C_2)} + \text{ 17 other terms with three contractions.} \end{aligned} \tag{W.296}$$

Thus, the time ordering at the l.h.s is replaced by 64 different terms at the r.h.s, provided that all contractions exist.

Fortunately, in QED there survive only a few contractions as we know from (W.285), namely

$$\begin{aligned} \wick{\c\phi (x) \c\phi^\dagger} (y) &= i \Delta_F (x - y) \\ \wick{\c\psi_\alpha (x) \c{\bar\psi}_\beta} (y) = -\wick{\c{\bar\psi}_\beta (y) \c\psi_\alpha} (x) &= i S_{F\alpha\beta} (x - y) \\ \wick{\c A^\mu (x) \c A^\nu} (y) &= i D_F^{\mu\nu} (x - y) . \end{aligned} \tag{W.297}$$

All other contractions vanish as e.g. $\wick{\c\psi(x) \c A^\mu}(y) = 0$ or $\wick{\c\psi(x) \c\psi}(y) = 0$. This fact reduces the number of terms considerably, as we will see in the next section.

W.10.4 Exercises and Solutions

1. Show (W.265).
   Solution: With

$$\mathcal{T} [A (x) B (y)] = \begin{cases} A (x) B (y) & \text{for } x^0 > y^0 \\ \pm B (y) A (x) & \phantom{\text{for }} y^0 > x^0 \end{cases} \tag{W.298}$$

   we arrive at

$$\wick{\c A(x) \c B}(y) = \begin{cases} A (x) B (y) - A (x) B (y) + \left[A^d (x) , B^c (y)\right]_\mp & \text{for } x^0 > y^0 \\ \pm B (y) A (x) - A (x) B (y) + \left[A^d (x) , B^c (y)\right]_\mp & \phantom{\text{for }} y^0 > x^0 \end{cases} . \tag{W.299}$$

The case $x^0 > y^0$ is clear; for $y^0 > x^0$ we have to calculate $\pm B(y) A(x) - A(x) B(y)$. It is

$$\begin{aligned} \pm B(y) A(x) - A(x) B(y) &= -[A(x) B(y) \mp B(y) A(x)] = -[A(x), B(y)]_{\mp} = \\ &= -\left[A^d(x) + A^c(x), B^d(y) + B^c(y)\right]_{\mp} = \\ &= -\left[A^d(x), B^d(y)\right]_{\mp} - \left[A^d(x), B^c(y)\right]_{\mp} - \left[A^c(x), B^d(y)\right]_{\mp} - \left[A^c(x), B^c(y)\right]_{\mp} = \\ &= -\left[A^d(x), B^c(y)\right]_{\mp} - \left[A^c(x), B^d(y)\right]_{\mp} \end{aligned} \tag{W.300}$$

since $\left[A^d(x), B^d(y)\right]_{\mp} = [A^c(x), B^c(y)]_{\mp} = 0$. It follows

$$\begin{aligned} &\pm B(y) A(x) - A(x) B(y) + \left[A^d(x), B^c(y)\right]_{\mp} = \\ &= -\left[A^d(x), B^c(y)\right]_{\mp} - \left[A^c(x), B^d(y)\right]_{\mp} + \left[A^d(x), B^c(y)\right]_{\mp} = . \\ &= -\left[A^c(x), B^d(y)\right]_{\mp} \end{aligned} \tag{W.301}$$

Collecting the results, we have

$$\overbrace{A(x)B(y)} = \begin{cases} \left[A^d(x), B^c(y)\right]_{\mp} & \text{for } x^0 > y^0 \\ -\left[A^c(x), B^d(y)\right]_{\mp} & \phantom{\text{for }} y^0 > x^0 \end{cases}. \tag{W.302}$$

2. Prove (W.278).
Solution: We have to show

$$\begin{aligned} &\theta\left(x^0 - y^0\right) \langle 0| \phi(x) \phi^\dagger(y) |0\rangle + \theta\left(y^0 - x^0\right) \langle 0| \phi^\dagger(y) \phi(x) |0\rangle) = \\ &= \theta\left(x^0 - y^0\right) \langle 0| \left[\phi^d(x), \phi^{\dagger c}(y)\right] |0\rangle + \theta\left(y^0 - x^0\right) \langle 0| \left[\phi^{\dagger d}(y), \phi^c(x)\right] |0\rangle \end{aligned} \tag{W.303}$$

i.e.,

$$\langle 0| \phi(x) \phi^\dagger(y) |0\rangle = \langle 0| \left[\phi^d(x), \phi^{\dagger c}(y)\right] |0\rangle \;\; ; \;\; \langle 0| \phi^\dagger(y) \phi(x) |0\rangle = \langle 0| \left[\phi^{\dagger d}(y), \phi^c(x)\right] |0\rangle . \tag{W.304}$$

We consider first $\langle 0| \phi(x) \phi^\dagger(y) |0\rangle$. With $\phi(x) = \phi^d(x) + \phi^c(x)$ and $\phi^\dagger(y) = \phi^{\dagger d}(y) + \phi^{\dagger c}(y)$ we have

$$\begin{aligned} \langle 0| \phi(x) \phi^\dagger(y) |0\rangle &= \langle 0| \phi^d(x) \phi^{\dagger d}(y) |0\rangle + \langle 0| \phi^d(x) \phi^{\dagger c}(y) |0\rangle \\ &+ \langle 0| \phi^c(x) \phi^{\dagger d}(y) |0\rangle + \langle 0| \phi^c(x) \phi^{\dagger c}(y) |0\rangle . \end{aligned} \tag{W.305}$$

A destruction (or annihilation) operator applied to the vacuum yields zero, $\phi^{\dagger d}(y) |0\rangle$; thus, the first and the third term on the r.h.s vanish. Likewise, $\langle 0| \phi^c(x)$ vanishes. So we have in a first step

$$\langle 0| \phi(x) \phi^\dagger(y) |0\rangle = \langle 0| \phi^d(x) \phi^{\dagger c}(y) |0\rangle . \tag{W.306}$$

On the r.h.s, we subtract $0 = \langle 0| \phi^{\dagger c}(y) \phi^d(x) |0\rangle$ and obtain

$$\langle 0|\,\phi(x)\,\phi^{\dagger}(y)\,|0\rangle = \langle 0|\,\phi^{d}(x)\,\phi^{\dagger c}(y)\,|0\rangle - \langle 0|\,\phi^{\dagger c}(y)\,\phi^{d}(x)\,|0\rangle = \langle 0|\left[\phi^{d}(x)\,,\phi^{\dagger c}(y)\right]|0\rangle\,. \tag{W.307}$$

In a similar way, we arrive at

$$\langle 0|\,\phi^{\dagger}(y)\,\phi(x)\,|0\rangle = \langle 0|\left[\phi^{\dagger d}(y)\,,\phi^{c}(x)\right]|0\rangle \tag{W.308}$$

which proves the assertion.

2. Prove (W.280).
   Solution: The Klein–Gordon free solution reads (cf. Appendix U, Vol. 1)

$$\phi^{\dagger}(y) = \frac{1}{(2\pi)^{3/2}}\int\frac{d^3p}{\sqrt{2E_{\mathbf{p}}}}\left[a_{\mathbf{p}}^{\dagger}e^{ipy} + a_{\mathbf{p}}e^{-ipy}\right]. \tag{W.309}$$

With $a_{\mathbf{p}}\,|0\rangle = 0$ and $a_{\mathbf{p}}^{\dagger}\,|0\rangle = |\mathbf{p}\rangle$ follows

$$\phi^{\dagger}(y)\,|0\rangle = \frac{1}{(2\pi)^{3/2}}\int\frac{d^3p}{\sqrt{2E_{\mathbf{p}}}}\,|\mathbf{p}\rangle\, e^{ipy}. \tag{W.310}$$

Taking the Hermitian adjoint $\left(\phi^{\dagger}(y)\,|0\rangle\right)^{\dagger} = \langle 0|\,\phi(y)$ brings (with $y \to x$ and $\mathbf{p} \to \mathbf{p}'$)

$$\langle 0|\,\phi(x) = \frac{1}{(2\pi)^{3/2}}\int\frac{d^3p'}{\sqrt{2E_{\mathbf{p}'}}}\left\langle \mathbf{p}'\right|e^{-ip'x}. \tag{W.311}$$

Multiplying the last two expressions leads to

$$\langle 0|\,\phi(x)\,\phi^{\dagger}(y)\,|0\rangle = \frac{1}{(2\pi)^{3}}\int\frac{d^3p\,d^3p'}{\sqrt{2E_{\mathbf{p}}}\sqrt{2E_{\mathbf{p}'}}}\left\langle \mathbf{p}'\right|\mathbf{p}\rangle\, e^{ip(y-x)}. \tag{W.312}$$

As is usual, we assume the momentum states as orthonormalized, i.e., $\left\langle \mathbf{p}'\right|\mathbf{p}\rangle = \delta\left(\mathbf{p}' - \mathbf{p}\right)$. Thus, we have

$$\langle 0|\,\phi(x)\,\phi^{\dagger}(y)\,|0\rangle = \frac{1}{(2\pi)^{3}}\int\frac{d^3p}{\sqrt{2E_{\mathbf{p}}}}e^{ip(y-x)}. \tag{W.313}$$

In a similar way, we have

$$\langle 0|\,\phi^{\dagger}(y)\,\phi(x)\,|0\rangle = \frac{1}{(2\pi)^{3}}\int\frac{d^3p}{\sqrt{2E_{\mathbf{p}}}}e^{ip(x-y)}. \tag{W.314}$$

Due to

$$i\Delta_F(x,y) = \langle 0|\,\mathcal{T}\left[\phi(x)\,\phi^\dagger(y)\right]|0\rangle =$$
$$= \theta\left(x^0 - y^0\right)\ \langle 0|\,\phi(x)\,\phi^\dagger(y)\,|0\rangle + \theta\left(y^0 - x^0\right)\ \langle 0|\,\phi^\dagger(y)\,\phi(x)\,|0\rangle \tag{W.315}$$

we arrive at the result[127]

$$i\Delta_F(x,y) = \frac{1}{(2\pi)^3}\int \frac{d^3p}{\sqrt{2E_\mathbf{p}}}\left[\theta\left(x^0 - y^0\right)\ e^{-ip(x-y)} + \theta\left(y^0 - x^0\right)\ e^{ip(x-y)}\right]. \tag{W.316}$$

W.11 *S*-Matrix, 2. Order, General

We now focus on the discussion of the second order term of the S-matrix $S^{(2)} = (-i)^2\,\frac{1}{2!}\int\int d^4x\, d^4y\, T\ [\mathcal{H}_I(x)\,\mathcal{H}_I(y)]$ with $\mathcal{H}_I(x) = q\mathcal{N}\left[\bar{\psi}\not{A}\psi\right]$. First we transform $\mathcal{T}\ [\mathcal{H}_I(x)\,\mathcal{H}_I(y)]$ by means of Wick's theorem. Then we give an overview of the possible physical processes. In the next section, we discuss two of them in more detail, namely Bhabha and Møller scattering.

Written out in full we have[128]

$$S^{(2)} = -\frac{q^2}{2}\int\int d^4x_1\, d^4x_2\ T\ \left[\mathcal{N}\left[\bar{\psi}(x_1)\,\not{A}(x_1)\,\psi(x_1)\right]\mathcal{N}\left[\bar{\psi}(x_2)\,\not{A}(x_2)\,\psi(x_2)\right]\right]. \tag{W.317}$$

On behalf of a more transparent notation, we use the abbreviation

$$\bar{\psi}(x_1)\,\not{A}(x_1)\,\psi(x_1) = \left(\bar{\psi}\not{A}\psi\right)_1 \tag{W.318}$$

and can write

$$S^{(2)} = -\frac{q^2}{2}\int\int d^4x_1\, d^4x_2\ T\ \left[\mathcal{N}\left[\left(\bar{\psi}\not{A}\psi\right)_1\right]\mathcal{N}\left[\left(\bar{\psi}\not{A}\psi\right)_2\right]\right]. \tag{W.319}$$

W.11.1 Applying Wick's Theorem

We keep in mind that only the following contractions exist (see the last section, (W.285)):

$$\begin{gathered}\overbrace{\psi_\alpha(x)\,\bar{\psi}_\beta}(y) = -\overbrace{\bar{\psi}_\beta(y)\,\psi_\alpha}(x) = iS_{F\alpha\beta}(x-y)\\ \overbrace{A^\mu(x)\,A^\nu}(y) = iD_F^{\mu\nu}(x-y)\,.\end{gathered} \tag{W.320}$$

[127]Since $\Delta(x,y)$ depends on the difference $x-y$, it is sometimes written $\Delta(x-y)$.

[128]Note that $x_1$and x_2 are dummy indices; interchanging them must leave the physics unchanged. The fact that either x_1 or x_2 can occur first is accounted for by time ordering.

All other contractions vanish as e.g. $\wick{\c{\psi}(x)\c{A}^\mu(y)} = 0$ or $\wick{\c{\psi}(x)\c{\psi}(y)} = 0$.

By means of Wick's theorem, we can replace the time ordering in (W.319) by the following sum of normal ordered strings:

$$\begin{aligned}
\mathcal{T}\ \left[\mathcal{N}\left[\left(\bar\psi A\!\!\!/\psi\right)_1\right]\mathcal{N}\left[\left(\bar\psi A\!\!\!/\psi\right)_2\right]\right] = \mathcal{N}\left[\left(\bar\psi A\!\!\!/\psi\right)_1\left(\bar\psi A\!\!\!/\psi\right)_2\right] + \\
+\mathcal{N}\left[\wick{(\bar\psi \c1{A\!\!\!/}\psi)_1(\bar\psi \c1{A\!\!\!/}\psi)_2}\right] + \mathcal{N}\left[\wick{(\c1{\bar\psi} A\!\!\!/\psi)_1(\bar\psi A\!\!\!/\c1\psi)_2}\right] + \mathcal{N}\left[\wick{(\bar\psi A\!\!\!/\c1\psi)_1(\c1{\bar\psi} A\!\!\!/\psi)_2}\right] + \\
+\mathcal{N}\left[\wick{(\c1{\bar\psi} \c2{A\!\!\!/}\psi)_1(\bar\psi \c2{A\!\!\!/}\c1\psi)_2}\right] + \mathcal{N}\left[\wick{(\bar\psi \c1{A\!\!\!/}\c2\psi)_1(\c2{\bar\psi} \c1{A\!\!\!/}\psi)_2}\right] + \mathcal{N}\left[\wick{(\c1{\bar\psi} A\!\!\!/\c2\psi)_1(\c2{\bar\psi} A\!\!\!/\c1\psi)_2}\right] + \\
+\mathcal{N}\left[\wick{(\c1{\bar\psi} \c2{A\!\!\!/}\c3\psi)_1(\c3{\bar\psi} \c2{A\!\!\!/}\c1\psi)_2}\right].
\end{aligned} \tag{W.321}$$

As we see, there is one term without, three terms with one, three terms with two, and one term with three contractions.

Remember that a contraction, i.e., a propagator contains no destruction or creation operators. Hence we can move it out from normal ordered strings and from matrix elements:

$$\langle f|\,\mathcal{N}\left[\wick{(\bar\psi \c1{A\!\!\!/}\psi)_1(\bar\psi \c1{A\!\!\!/}\psi)_2}\right]|i\rangle = iD_F^{\mu\nu}\left(x_1 - x_2\right)\langle f|\,\mathcal{N}\left[\left(\bar\psi\gamma_\mu\psi\right)_1\left(\bar\psi\gamma_\nu\psi\right)_2\right]|i\rangle\,. \tag{W.322}$$

W.11.2 *Physical Interpretation*

Without going into details for the present, we give now an overview of the physical meaning of the different terms in (W.321). In the following, electron/positron means electron *or* positron, and electron-positron stands for electron *and* positron.[129] The Feynman diagrams show typical processes.

W.11.2.1 No Contraction

The term $\mathcal{N}\left[\left(\bar\psi A\!\!\!/\psi\right)_1\left(\bar\psi A\!\!\!/\psi\right)_2\right]$ represents two independent processes of first order, as we have discussed them above in section 'S-matrix,first order'. Since all these processes are not real physical processes, we ignore the term.

[129] The considerations also apply to muons and taus and their antiparticles, of course.

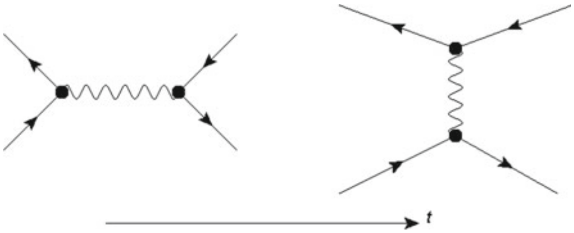


Fig. W.7 Bhabha scattering type 1 (annihilation) and type 2 (direct scattering)

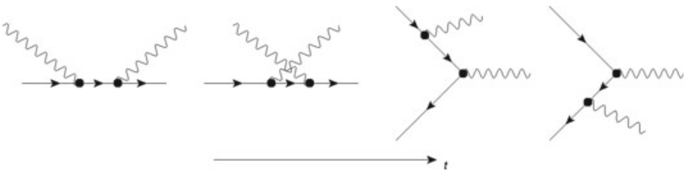


Fig. W.8 Compton scattering and pair annihilation

W.11.2.2 One Contraction

Photon propagator In this case, we have

$$S^{(2)} = -\frac{q^2}{2} \int \int \mathrm{d}^4 x_1 \, \mathrm{d}^4 x_2 \, \mathcal{N} \left[(\bar{\psi} \not{A} \psi)_1 (\bar{\psi} \not{A} \psi)_2 \right]. \tag{W.323}$$

The contraction is given by the Feynman propagator $iD_F^{\mu\nu}(x_1 - x_2)$. Thus, the only creation or destruction operators are found in $\bar{\psi}$ and ψ. This means, that the only initial states that can be destroyed are electron/positron states, and the same holds true for the creation in the final states. This type of interaction is called four external lepton interaction, as e.g. scattering of electrons and positrons. As an example, we show Bhabha scattering in Fig. W.7, i.e., scattering of electron and positron. Below, we consider this case in more detail.

A remark concerning the naming. The in- and outgoing particles are called *external particles*, all other particles are called *internal particles*.

Fermion propagator There are two terms with a fermion propagator which can shown to be equal. So we have in this case

$$S^{(2)} = -q^2 \int \int \mathrm{d}^4 x_1 \, \mathrm{d}^4 x_2 \, \mathcal{N} \left[(\bar{\psi} \not{A} \psi)_1 (\bar{\psi} \not{A} \psi)_2 \right]. \tag{W.324}$$

The contraction is given by the fermion propagator $iS_{F\alpha\beta}(x_1 - x_2)$. The states which can be destroyed or created comprise a photon and a electron/positron. Typical processes are Compton scattering or pair annihilation, see Fig. W.8.

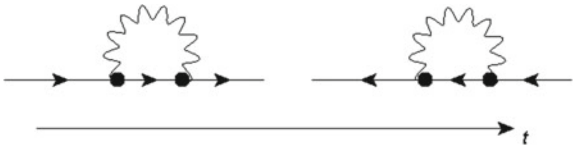


Fig. W.9 Electron and positron loop

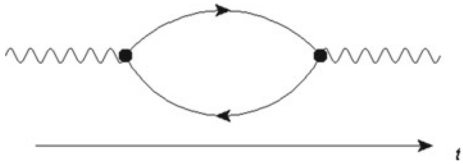


Fig. W.10 Photon closed loop

W.11.2.3 Two Contractions

One fermion and one photon propagator There are two such terms in (W.321) which again can be shown to be equal. So we have in this case

$$S^{(2)} = -q^2 \int \int \mathrm{d}^4x_1\, \mathrm{d}^4x_2\, \mathcal{N}\left[(\bar{\psi} \not{A} \psi)_1 (\bar{\psi} \not{A} \psi)_2 \right]. \tag{W.325}$$

This represents destruction of an electron/positron, followed by the propagation of both the electron/positron and the photon, and finally the creation of an electron/positron. These processes are called electron (or positron) closed loop. The name stems from the pictorial representation, see Fig. W.9. Another name is self-energy diagrams.

Two fermion propagators Here we have

$$S^{(2)} = -\frac{q^2}{2} \int \int \mathrm{d}^4x_1\, \mathrm{d}^4x_2\, \mathcal{N}\left[(\bar{\psi} \not{A} \psi)_1 (\bar{\psi} \not{A} \psi)_2 \right]. \tag{W.326}$$

The two propagators are of the fermion propagator type, $iS_{F\alpha\beta}(x_1 - x_2)$. There is a real incoming and outgoing photon; these two photons are connected by a virtual electron-positron pair, see Fig. W.10. The process is called photon closed loop or vacuum closed loop (since ‘in between’ the vacuum splits up into a negative and a positive charged particle).

W.11.2.4 Three Contractions

This term is given by

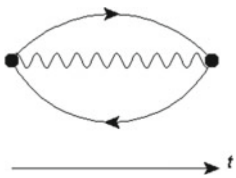


Fig. W.11 Vacuum bubble

$$S^{(2)} = -\frac{q^2}{2}\int\int \mathrm{d}^4x_1\ \mathrm{d}^4x_2\ \mathcal{N}\left[(\bar{\psi}\not{A}\psi)_1(\bar{\psi}\not{A}\psi)_2\right]. \tag{W.327}$$

There are only propagators, no annihilation and creation operators, i.e., only internal and no external lines. We can visualize the term as a process which starts and ends in the vacuum and splits up 'in between' in a virtual photon and a virtual electron-positron pair. The process is called vacuum bubble, see Fig. W.11.

W.12 *S*-Matrix, 2. Order, 4 Lepton Scattering

As a worked out example, we consider in this section in some detail the case 'one photon propagator', i.e., four lepton scattering. We determine and discuss explicitly transition amplitudes for two physically different processes, namely Bhabha and Møller scattering. By these examples, we illustrate the one-to-one correspondence between transition amplitudes and Feynman diagrams. Finally, we consider the connection between transition amplitude and the differential scattering cross section.

We start with (W.323)

$$S^{(2)} = -\frac{q^2}{2}\int\int \mathrm{d}^4x_1\ \mathrm{d}^4x_2\ \mathcal{N}\left[(\bar{\psi}\not{A}\psi)_1(\bar{\psi}\not{A}\psi)_2\right]. \tag{W.328}$$

The contraction is given by the Feynman propagator $iD_F^{\mu\nu}(x_1 - x_2)$. Thus, we can write

$$S^{(2)} = -\frac{q^2}{2}\int\int \mathrm{d}^4x_1\ \mathrm{d}^4x_2\ \mathcal{N}\left[\left(\bar{\psi}\gamma_\mu\psi\right)_1\left(\bar{\psi}\gamma_\nu\psi\right)_2\right] iD_F^{\mu\nu}(x_1 - x_2)\,. \tag{W.329}$$

As initial and final state we can choose electrons and positrons (of course, myons and tauons and their antiparticles would also be possible), whence the name *four external lepton interaction*. As a worked out example, we consider the case where initial and final states are an electron/positron pair:

$$|i\rangle = \left|e^-(\mathbf{p}_1 r_1)\right\rangle\left|e^+(\mathbf{p}_2 r_2)\right\rangle\ ;\ |f\rangle = \left|e^-\left(\mathbf{p}_1' r_1'\right)\right\rangle\left|e^+\left(\mathbf{p}_2' r_2'\right)\right\rangle. \tag{W.330}$$

This type of scattering (particle plus antiparticle) is called *Bhabha*[130] *scattering*. After this example, we transfer the results with the necessary modifications to describe the scattering of particle/particle pairs, e.g. two electrons, called *Møller*[131] *scattering*.

The purpose of the following considerations is twofold. First, we want to show step by step how to calculate transition amplitudes which is done for Bhabha scattering. Second, we will see that the Feynman diagrams (together with Feynman rules, see Table W.5) carry the same information as the mathematical formulation. This connection will be applied in case of Møller scattering, where we formulate the Feynman amplitude directly from the Feynman diagrams - and, hopefully, convince the reader that these diagrams are very functional and make life considerably easier.

W.12.1 Bhabha Scattering, $e^+e^- \rightarrow e^+e^-$

We now start the calculation of $S^{(2)}$ for this process.[132] The decomposition of $\mathcal{N}\left[\left(\bar{\psi}\gamma_\mu\psi\right)_1\left(\bar{\psi}\gamma_\nu\psi\right)_2\right]$ in (W.329) with respect to creation and deletion operators gives 16 terms. But we do not need to write them all down since only those will contribute which destroy incoming and create outgoing pairs of e^-/e^+.

Note that there are two possibilities in which in- and outgoing particles have the same individual momenta and spins. Either at e.g. x_2 both the e^- and the e^+ are destroyed, a photon runs from x_2 to x_1, and at x_1 both the e^- and the e^+ are created (plus the process with reversed roles of x_2 and x_1). In other words: we have first a pair annihilation producing a (virtual) photon, followed by pair production. We call this process Bhabha scattering type 1 or annihilation scattering; the transition amplitude is written as $S_{B1}^{(2)}$. Or the incoming e^- is destroyed at e.g. x_2 and the outgoing e^- is created at x_2; a photon starts at x_2 and ends in x_1 where the incoming e^+ is destroyed and again created (plus the process with reversed roles of x_2 and x_1). In other words: the two particles are directly scattered whereby the interaction is mediated by a (virtual) photon. We call this process Bhabha scattering type 2 or direct scattering; the transition amplitude is written as $S_{B2}^{(2)}$. The Feynman diagrams with the corresponding momenta are found in Fig. W.12. We repeat that the process with reserved role of x_2 and x_1 is also possible. Time ordering ensures the right chronological order of x_1 and x_2.

Note that we cannot distinguish the two processes by measurement. Indeed, we can only measure the incoming and outgoing particles, and hence we do not know which one of the two processes may have occurred. Consequently, quantum mechanics says

[130] Bhabha, Homi Jehangir; 1909–1966, Indian physicist.

[131] Møller, Christian; 1904–1980, Danish physicist.

[132] Note that Bhabha scattering is not only interesting from a theoretical point of view. Luminosity, i.e., the number of collisions per time and area, is a measure for the performance of a particle accelerator. In many colliders, luminosity is determined using Bhabha scattering at small angles.

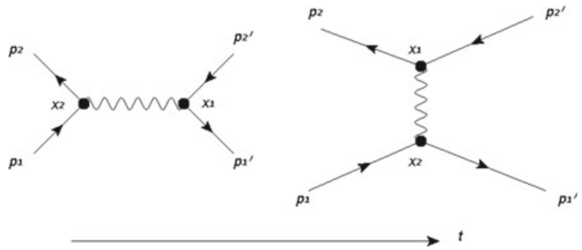


Fig. W.12 Bhabha scattering. Left annihilation (type 1), right direct scattering (type 2)

that we have to add the two amplitudes which means we have interference effects between the two possibilities.

In the discussion of the physical processes described by the first order S -matrix $S^{(1)}$ we have seen that one-vertex processes are not 'real'. However, they are the building blocks of real higher order processes. For instance, Bhabha scattering type 1 can be interpreted as composed of first order pair annihilation ($e^+ + e^- \rightarrow \gamma$), followed by first order pair production ($\gamma \rightarrow e^+ + e^-$). Incoming and outgoing particles are real whereas the photon is virtual.

W.12.1.1 Bhabha Scattering Type 1

Warning: Be prepared for a lot of algebra! Calculating a transition amplitude step by step takes time and place. All the nicer then to work with Feynman diagrams.

For x_2 the only combination which destroys $\left|e^-(\mathbf{p}_1 r_1)\right\rangle$ and $\left|e^+(\mathbf{p}_2 r_2)\right\rangle$ is given by $\left(\bar{\psi}^d \gamma_\nu \psi^d\right)_2$. The propagator creates a virtual photon at x_2 and propagates it to x_1 where it is destroyed. Then the outgoing pair of e^-/e^+ is created at x_1 by $\left(\bar{\psi}^c \gamma_\mu \psi^c\right)_1$.

The same consideration holds with reversed roles of x_1 and x_2. Thus, we have for $S_{B1}^{(2)} = \left\langle e^-\left(\mathbf{p}_1' r_1'\right) e^+\left(\mathbf{p}_2' r_2'\right)\right| S^{(2)} \left|e^-(\mathbf{p}_1 r_1) e^+(\mathbf{p}_2 r_2)\right\rangle$ the expression

$$\begin{aligned} S_{B1}^{(2)} &= -\tfrac{q^2}{2} \left\langle e^-\left(\mathbf{p}_1' r_1'\right) e^+\left(\mathbf{p}_2' r_2'\right)\right| \int\int \mathrm{d}^4x_1\, \mathrm{d}^4x_2\, iD_F^{\mu\nu}(x_1 - x_2) \cdot \\ &\cdot \left\{\mathcal{N}\left[\left(\bar{\psi}^c \gamma_\mu \psi^c\right)_1 \left(\bar{\psi}^d \gamma_\nu \psi^d\right)_2\right] + \mathcal{N}\left[\left(\bar{\psi}^d \gamma_\nu \psi^d\right)_1 \left(\bar{\psi}^c \gamma_\mu \psi^c\right)_2\right]\right\} \left|e^-(\mathbf{p}_1 r_1) e^+(\mathbf{p}_2 r_2)\right\rangle. \end{aligned} \tag{W.331}$$

Normal ordering The first term in the curled brackets is already normal ordered. In the second term, we switch ψ^d first with $\bar{\psi}^c$ which gives a minus sign, and after that with ψ^c which introduces a second minus sign; thus, we have no change altogether. The same holds for switching $\bar{\psi}^d$, and the result reads

$$\begin{aligned} S_{B1}^{(2)} &= -\tfrac{q^2}{2} \left\langle e^-\left(\mathbf{p}_1' r_1'\right) e^+\left(\mathbf{p}_2' r_2'\right)\right| \int\int \mathrm{d}^4x_1\, \mathrm{d}^4x_2 \cdot \\ &\cdot \left\{\left(\bar{\psi}^c \gamma_\mu \psi^c\right)_1 \left(\bar{\psi}^d \gamma_\nu \psi^d\right)_2 + \left(\bar{\psi}^c \gamma_\mu \psi^c\right)_2 \left(\bar{\psi}^d \gamma_\nu \psi^d\right)_1\right\} iD_F^{\mu\nu}(x_1 - x_2) \left|e^-(\mathbf{p}_1 r_1) e^+(\mathbf{p}_2 r_2)\right\rangle. \end{aligned} \tag{W.332}$$

For the second term, we switch the (dummy) integration variables[133] and arrive at

[133]Remember $D_F^{\mu\nu}(x_1 - x_2) = D_F^{\mu\nu}(x_2 - x_1)$.

$$\begin{aligned} S_{B1}^{(2)} &= -q^2 \left\langle e^- \left(\mathbf{p}_1' r_1'\right) e^+ \left(\mathbf{p}_2' r_2'\right)\right| \int \int \mathrm{d}^4 x_1\ \mathrm{d}^4 x_2 \cdot \\ &\cdot \left(\bar{\psi}^c \gamma_\mu \psi^c\right)_1 iD_F^{\mu\nu}\left(x_1 - x_2\right) \left(\bar{\psi}^d \gamma_\nu \psi^d\right)_2 \left| e^- \left(\mathbf{p}_1 r_1\right) e^+ \left(\mathbf{p}_2 r_2\right)\right\rangle . \end{aligned} \tag{W.333}$$

Here we have placed the propagator (which is just a c−number) between the field operators in order to display the physical process.

Inserting the field operators, evaluation of the brackets We now insert the field operators (discrete version) given by (see section 'Quantization of free fields, Dirac')[134]

$$\begin{aligned} \psi(x) &= \sum_{\mathbf{p},r} \sqrt{\frac{m}{VE_\mathbf{p}}} \left(b_r(\mathbf{p})\, u_r(\mathbf{p})\, e^{-ipx} + \mathrm{d}_r^\dagger(\mathbf{p})\, w_r(\mathbf{p})\, e^{ipx}\right) \\ \bar{\psi}(x) &= \sum_{\mathbf{p},r} \sqrt{\frac{m}{VE_\mathbf{p}}} \left(b_r^\dagger(\mathbf{p})\, \bar{u}_r(\mathbf{p})\, e^{ipx} + \mathrm{d}_r(\mathbf{p})\, \bar{w}_r(\mathbf{p})\, e^{-ipx}\right). \end{aligned} \tag{W.334}$$

As we know from the discussion of the first order S-matrix (see section 'S-matrix, first order'), applying the field operator $\left(\bar{\psi}^d \gamma_\nu \psi^d\right)_2$ to the initial state picks out the elements with the same quantum numbers,[135] i.e.,

$$\left(\bar{\psi}^d \gamma_\nu \psi^d\right)_2 \left| e^- \left(\mathbf{p}_1 r_1\right) e^+ \left(\mathbf{p}_2 r_2\right)\right\rangle = \sqrt{\frac{m}{VE_{\mathbf{p}_2}}} \bar{w}_{r_2}\left(\mathbf{p}_2\right) e^{-ip_2 x_2} \gamma_\nu \sqrt{\frac{m}{VE_{\mathbf{p}_1}}} u_{r_1}\left(\mathbf{p}_1\right) e^{-ip_1 x_2} \left|0\right\rangle . \tag{W.335}$$

In the same way, we have

$$\left\langle e^- \left(\mathbf{p}_1' r_1'\right) e^+ \left(\mathbf{p}_2' r_2'\right)\right| \left(\bar{\psi}^c \gamma_\mu \psi^c\right)_1 = \left\langle 0\right| \sqrt{\frac{m}{VE_{\mathbf{p}_1'}}} \bar{u}_{r_1'}\left(\mathbf{p}_1'\right) e^{ip_1' x_1} \gamma_\mu \sqrt{\frac{m}{VE_{\mathbf{p}_2'}}} w_{r_2'}\left(\mathbf{p}_2'\right) e^{ip_2' x_1} . \tag{W.336}$$

Combining the results yields

$$\begin{aligned} S_{B1}^{(2)} &= -q^2 \left\langle 0\right| \int \int \mathrm{d}^4 x_1\ \mathrm{d}^4 x_2 \sqrt{\frac{m}{VE_{\mathbf{p}_1'}}} \bar{u}_{r_1'}\left(\mathbf{p}_1'\right) e^{ip_1' x_1} \gamma_\mu \sqrt{\frac{m}{VE_{\mathbf{p}_2'}}} \cdot \\ &\cdot w_{r_2'}\left(\mathbf{p}_2'\right) e^{ip_2' x_1} iD_F^{\mu\nu}\left(x_1 - x_2\right) \sqrt{\frac{m}{VE_{\mathbf{p}_2}}} \bar{w}_{r_2}\left(\mathbf{p}_2\right) e^{-ip_2 x_2} \gamma_\nu \sqrt{\frac{m}{VE_{\mathbf{p}_1}}} u_{r_1}\left(\mathbf{p}_1\right) e^{-ip_1 x_2} \left|0\right\rangle . \end{aligned} \tag{W.337}$$

As we see, the term Θ between the bra-ket $\left\langle 0\right| \Theta \left|0\right\rangle$ contains neither creation nor destruction operators. Thus, we can write $\left\langle 0\right| \Theta \left|0\right\rangle = \Theta \left\langle 0\right|\left. 0\right\rangle = \Theta$ due to $\left\langle 0\right|\left. 0\right\rangle = 1$. It follows

$$\begin{aligned} S_{B1}^{(2)} &= -q^2 \int \int \mathrm{d}^4 x_1\ \mathrm{d}^4 x_2 \sqrt{\frac{m}{VE_{\mathbf{p}_1'}}} \sqrt{\frac{m}{VE_{\mathbf{p}_2'}}} \sqrt{\frac{m}{VE_{\mathbf{p}_2}}} \sqrt{\frac{m}{VE_{\mathbf{p}_1}}} \cdot \\ &\cdot \bar{u}_{r_1'}\left(\mathbf{p}_1'\right) \gamma_\mu w_{r_2'}\left(\mathbf{p}_2'\right) \bar{w}_{r_2}\left(\mathbf{p}_2\right) \gamma_\nu u_{r_1}\left(\mathbf{p}_1\right) \cdot iD_F^{\mu\nu}\left(x_1 - x_2\right) e^{ip_1' x_1} e^{ip_2' x_1} e^{-ip_2 x_2} e^{-ip_1 x_2} . \end{aligned} \tag{W.338}$$

Calculating the x-integral What remains is to calculate the x_1/x_2- integrals X

---

[134]For the continuous version replace the sum by an integral and V by $(2\pi)^3$.

[135]Remember that $b_r^\dagger(\mathbf{p})$ / $b_r(\mathbf{p})$ creates /deletes an electron and $\mathrm{d}_r^\dagger(\mathbf{p})$ / $\mathrm{d}_r(\mathbf{p})$ creates /deletes a positron.

$$X = \int \int \mathrm{d}^4 x_1 \, \mathrm{d}^4 x_2 \, D_F^{\mu\nu} (x_1 - x_2) \, e^{iP_1' x_1} e^{-iPx_2} \;\; ; P = p_1 + p_2 \; ; \; P' = p_1' + p_2'. \tag{W.339}$$

The result reads (see exercises)

$$X = D_F^{\mu\nu} (P) \, (2\pi)^4 \, \delta^{(4)} \left(P - P'\right) \tag{W.340}$$

where $D_F^{\mu\nu} (P)$ is the Fourier transform of the propagator, i.e.,

$$D_F^{\mu\nu} (P) = -g^{\mu\nu} \frac{1}{P^2 + i\varepsilon}. \tag{W.341}$$

Final result for $S_{B1}^{(2)}$ Thus, we have for the S-matrix the result

$$\begin{array}{c} S_{B1}^{(2)} = \sqrt{\frac{m}{VE_{\mathbf{p}_1'}}} \sqrt{\frac{m}{VE_{\mathbf{p}_2'}}} \sqrt{\frac{m}{VE_{\mathbf{p}_2}}} \sqrt{\frac{m}{VE_{\mathbf{p}_1}}} \cdot (2\pi)^4 \, \delta^{(4)} \left(p_1 + p_2 - p_1' - p_2'\right) \cdot \\ \cdot \left\{ \left(-q^2\right) \bar{u}_{r_1'} \left(\mathbf{p}_1'\right) \gamma_\mu w_{r_2'} \left(\mathbf{p}_2'\right) i D_F^{\mu\nu} \left(p_1 + p_2\right) \bar{w}_{r_2} \left(\mathbf{p}_2\right) \gamma_\nu u_{r_1} \left(\mathbf{p}_1\right) \right\}. \end{array} \tag{W.342}$$

The term in curly brackets is called *Feynman amplitude* and noted as $\mathcal{M}_{B1}^{(2)}$. Thus, we can write the S-matrix for Bhabha scattering type 1 as

$$S_{B1}^{(2)} = \left(\prod_{\mathbf{p}}^{\text{external fermions}} \sqrt{\frac{m}{VE_{\mathbf{p}}}} \right) (2\pi)^4 \, \delta^{(4)} \left(p_2 + p_1 - \left(p_2' + p_1'\right)\right) \mathcal{M}_{B1}^{(2)} \tag{W.343}$$

where

$$\mathcal{M}_{B1}^{(2)} = \left(-q^2\right) \bar{u}_{r_1'} \left(\mathbf{p}_1'\right) \gamma_\mu w_{r_2'} \left(\mathbf{p}_2'\right) i D_F^{\mu\nu} \left(p_1 + p_2\right) \bar{w}_{r_2} \left(\mathbf{p}_2\right) \gamma_\nu u_{r_1} \left(\mathbf{p}_1\right). \tag{W.344}$$

Let us concentrate on this Feynman amplitude. From right to left, i.e., in the direction of time, we can 'read' $\mathcal{M}_{B1}^{(2)}$) as follows: $u_{r_1} (\mathbf{p}_1)$ destroys an electron with quantum numbers $(r_1, \mathbf{p}_1)$, $\bar{w}_{r_2} (\mathbf{p}_2)$ destroys a positron $(r_2, \mathbf{p}_2)$. A virtual photon with $k = p_1 + p_2$ is the link to $w_{r_2'} \left(\mathbf{p}_2'\right)$, creating a positron with $(r_2', \mathbf{p}_2')$, and $\bar{u}_{r_1'} \left(\mathbf{p}_1'\right)$, creating an electron with $(r_1', \mathbf{p}_1')$. This order is precisely reflected in the corresponding Feynman diagram (W.12). Following the arrows, we have: At the left, we have an incoming electron $u_{r_1} (\mathbf{p}_1)$, then a vertex which contributes a factor $-iq\gamma_\nu$, followed by an incoming positron $\bar{w}_{r_2} (\mathbf{p}_2)$. A virtual photon with momentum $p_1 + p_2$ (i.e., $iD_F^{\mu\nu} (p_1 + p_2)$) is the connection to the outgoing channel at the right. Here we have a positron with $w_{r_2'} \left(\mathbf{p}_2'\right)$, a vertex with $-iq\gamma_\mu$ and an electron with $\bar{u}_{r_1'} \left(\mathbf{p}_1'\right)$.

In other words, we could have saved the last two pages of calculation, simply by reading off the information provided by the Feynman diagram. We see that the spinor factors (γ matrices, fermionic propagators, 4-spinors) in (W.344) occur in exactly the same order as following in Fig. W.12 the fermion lines in the direction of the arrows through the vertices. Indeed, between Feynman amplitude and Feynman diagram there is a one-to-one relation which evidently can save a lot of computing.

To test these statements with another (and last) example, we consider now Bhabha scattering type 2.

W.12.1.2 Bhabha Scattering Type 2

Arguing along the same lines as for $S_{B1}^{(2)}$, given by (W.331), we can write for $S_{B2}^{(2)}$

$$S_{B2}^{(2)} = -\frac{q^2}{2} \left\langle e^- \left(\mathbf{p}_1' r_1'\right), e^+ \left(\mathbf{p}_2' r_2'\right)\right| \int \int \mathrm{d}^4 x_1 \, \mathrm{d}^4 x_2 \cdot \\ \cdot \left\{ \mathcal{N}\left[\left(\bar{\psi}^d \gamma_\mu \psi^c\right)_1 \left(\bar{\psi}^c \gamma_\nu \psi^d\right)_2\right] + \mathcal{N}\left[\left(\bar{\psi}^c \gamma_\mu \psi^d\right)_1 \left(\bar{\psi}^d \gamma_\nu \psi^c\right)_2\right]\right\} iD_F^{\mu\nu}\left(x_1 - x_2\right) \left|e^- \left(\mathbf{p}_1 r_1\right), e^+ \left(\mathbf{p}_2 r_2\right)\right\rangle . \tag{W.345}$$

Normal ordering Switching the integration variables x_1 and x_2 reveals that the two expressions in the curly brackets are identical. However, normal ordering these terms is a little bit more complicated as in the type 1 case. This is due to the mixed character of the brackets $\left(\bar{\psi}^d \gamma_\mu \psi^c\right)_1$ and $\left(\bar{\psi}^c \gamma_\nu \psi^d\right)_2$, where deletion and creation operator are found in one bracket. Thus, we can not simply interchange them but have to take into account the spinor indices. This means that we write[136]

$$S_{B2}^{(2)} = -q^2 \left\langle e^- \left(\mathbf{p}_1' r_1'\right), e^+ \left(\mathbf{p}_2' r_2'\right)\right| \int \int \mathrm{d}^4 x_1 \, \mathrm{d}^4 x_2 \cdot \\ \cdot \left\{ \mathcal{N}\left[\left(\bar{\psi}_\alpha^d \gamma_{\mu\alpha\beta} \psi_\beta^c\right)_1 \left(\bar{\psi}_\delta^c \gamma_{\nu\delta\eta} \psi_\eta^d\right)_2\right]\right\} iD_F^{\mu\nu}\left(x_1 - x_2\right) \left|e^- \left(\mathbf{p}_1 r_1\right), e^+ \left(\mathbf{p}_2 r_2\right)\right\rangle \tag{W.346}$$

where we use summation convention and $\gamma_{\mu\alpha\beta}$ is the element $(\alpha\beta)$ of γ_μ. Now we can normal order:

$$\mathcal{N}\left[\left(\bar{\psi}_\alpha^d \gamma_{\mu\alpha\beta} \psi_\beta^c\right)_1 \left(\bar{\psi}_\delta^c \gamma_{\nu\delta\eta} \psi_\eta^d\right)_2\right] = \mathcal{N}\left[\left(\bar{\psi}_\delta^c\right)_2 \left(\bar{\psi}_\alpha^d \gamma_{\mu\alpha\beta} \psi_\beta^c\right)_1 \left(\gamma_{\nu\delta\eta} \psi_\eta^d\right)_2\right] = \\ = \mathcal{N}\left[-\left(\bar{\psi}_\delta^c\right)_2 \left(\gamma_{\mu\alpha\beta} \psi_\beta^c\right)_1 \left(\bar{\psi}_\alpha^d\right)_1 \left(\gamma_{\nu\delta\eta} \psi_\eta^d\right)_2\right]. \tag{W.347}$$

The first ordering leaves the sign unchanged since the arguments differ (x_1 and x_2), the second ordering changes the sign, since the arguments are equal (both x_1). In this way we arrive at

$$S_{B2}^{(2)} = q^2 \left\langle e^- \left(\mathbf{p}_1' r_1'\right), e^+ \left(\mathbf{p}_2' r_2'\right)\right| \int \int \mathrm{d}^4 x_1 \, \mathrm{d}^4 x_2 \cdot \\ \cdot \left\{ \left(\bar{\psi}_\delta^c\right)_2 \gamma_{\mu\alpha\beta} \left(\psi_\beta^c\right)_1 \left(\bar{\psi}_\alpha^d\right)_1 \gamma_{\nu\delta\eta} \left(\psi_\eta^d\right)_2 \right\} iD_F^{\mu\nu}\left(x_1 - x_2\right) \left|e^- \left(\mathbf{p}_1 r_1\right), e^+ \left(\mathbf{p}_2 r_2\right)\right\rangle . \tag{W.348}$$

Note we have the same normal order as in type 1 scattering: the order from the right side moving leftward is e^- deletion, e^+ deletion, e^+ creation, and e^- creation, as it should be.

Inserting the field operators, evaluation of the brackets Inserting the field operators leads to

[136]We use the Einstein summation convention.

$$S_{B2}^{(2)} = q^2 \left\langle e^- \left(\mathbf{p}_1' r_1'\right), e^+ \left(\mathbf{p}_2' r_2'\right)\right| \int\int \mathrm{d}^4x_1\, \mathrm{d}^4x_2 \cdot \\ \cdot \left\{\left(\bar{\psi}_\delta^c\right)_2 \gamma_{\mu\alpha\beta} \left(\psi_\beta^c\right)_1 \left(\bar{\psi}_\alpha^d\right)_1 \gamma_{\nu\delta\eta} \left(\psi_\eta^d\right)_2\right\} iD_F^{\mu\nu}\left(x_1 - x_2\right) \left|e^- \left(\mathbf{p}_1 r_1\right), e^+ \left(\mathbf{p}_2 r_2\right)\right\rangle. \tag{W.349}$$

For a short auxiliary calculation, we neglect for a moment the integrals and the propagator,[137] thus defining an short-lived variable K by

$$K = \left\langle e^- \left(\mathbf{p}_1' r_1'\right), e^+ \left(\mathbf{p}_2' r_2'\right)\right| \left\{\left(\bar{\psi}_\delta^c\right)_2 \gamma_{\mu\alpha\beta} \left(\psi_\beta^c\right)_1 \left(\bar{\psi}_\alpha^d\right)_1 \gamma_{\nu\delta\eta} \left(\psi_\eta^d\right)_2\right\} \left|e^- \left(\mathbf{p}_1 r_1\right), e^+ \left(\mathbf{p}_2 r_2\right)\right\rangle. \tag{W.350}$$

Taking into account the above mentioned 'matching rules' (cf. section 'S -matrix, first order'), the evaluation of this term gives

$$K = \langle 0| \left\{ \begin{array}{c} \sqrt{\frac{m}{VE_{\mathbf{p}_1'}}} \bar{u}_{r_{1\delta}'}\left(\mathbf{p}_1'\right) e_2^{ip_1'x_2} \gamma_{\mu\alpha\beta} \sqrt{\frac{m}{VE_{\mathbf{p}_2'}}} w_{r_2'\beta}\left(\mathbf{p}_2'\right) e^{ip_2'x_1} \cdot \\ \cdot \sqrt{\frac{m}{VE_{\mathbf{p}_2}}} \bar{w}_{r_2\alpha}\left(\mathbf{p}_2\right) e_1^{-ip_2x_1} \gamma_{\nu\delta\eta} \sqrt{\frac{m}{VE_{\mathbf{p}_1}}} u_{r_1\eta}\left(\mathbf{p}_1\right) e^{-ip_1x_2} \end{array} \right\} |0\rangle. \tag{W.351}$$

Again, the term inside the bra–ket $\langle 0|\, \Theta\, |0\rangle$ contains neither creation nor destruction operators whence due to $\langle 0\, |0\rangle = 1$ follows $\langle 0|\, \Theta\, |0\rangle = \Theta \langle 0\, |0\rangle = \Theta$.

Now we rearrange the indexed terms of this expression. Note that $\bar{u}_{r_{1\delta}'}$, $\gamma_{\mu\alpha\beta}$ etc. are single elements of the vectors and matrices, i.e., scalars which can be moved freely. We have[138]

$$\bar{u}_{r_{1\delta}'}\left(\mathbf{p}_1'\right) \gamma_{\mu\alpha\beta} w_{r_2'\beta}\left(\mathbf{p}_2'\right) \bar{w}_{r_2\alpha}\left(\mathbf{p}_2\right) \gamma_{\nu\delta\eta} u_{r_1\eta}\left(\mathbf{p}_1\right) = \\ = \bar{w}_{r_2\alpha}\left(\mathbf{p}_2\right) \gamma_{\mu\alpha\beta} w_{r_2'\beta}\left(\mathbf{p}_2'\right) \bar{u}_{r_{1\delta}'}\left(\mathbf{p}_1'\right) \gamma_{\nu\delta\eta} u_{r_1\eta}\left(\mathbf{p}_1\right) = \bar{w}_{r_2}\left(\mathbf{p}_2\right) \gamma_\mu w_{r_2'}\left(\mathbf{p}_2'\right) \cdot \bar{u}_{r_1'}\left(\mathbf{p}_1'\right) \gamma_\nu u_{r_1}\left(\mathbf{p}_1\right). \tag{W.352}$$

Inserting these partial results into (W.349) yields

$$S_{B2}^{(2)} = q^2 \int\int \mathrm{d}^4x_1\, \mathrm{d}^4x_2 \cdot \sqrt{\frac{m}{VE_{\mathbf{p}_1'}}} \sqrt{\frac{m}{VE_{\mathbf{p}_2'}}} \sqrt{\frac{m}{VE_{\mathbf{p}_2}}} \sqrt{\frac{m}{VE_{\mathbf{p}_1}}} \cdot \\ \cdot \left\{\bar{w}_{r_2}\left(\mathbf{p}_2\right) \gamma_\mu w_{r_2'}\left(\mathbf{p}_2'\right) \cdot \bar{u}_{r_1'}\left(\mathbf{p}_1'\right) \gamma_\nu u_{r_1}\left(\mathbf{p}_1\right)\right\} e^{i\left(p_2'-p_2\right)x_1} e^{i\left(p_1'-p_1\right)x_2} iD_F^{\mu\nu}\left(x_1 - x_2\right). \tag{W.353}$$

Calculating the x-integral, final result for $S_{B2}^{(2)}$ Concerning the x-integration we can use the result just derived above, and obtain

$$\int\int \mathrm{d}^4x_1\, \mathrm{d}^4x_2\, e^{i\left(p_2'-p_2\right)x_1} e^{i\left(p_1'-p_1\right)x_2} iD_F^{\mu\nu}\left(x_1 - x_2\right) = \\ = iD_F^{\mu\nu}\left(p_2' - p_2\right) (2\pi)^4\, \delta^{(4)}\left(p_1' - p_1 + \left(p_2' - p_2\right)\right). \tag{W.354}$$

Thus, we arrive finally at

$$S_{B2}^{(2)} = q^2 \sqrt{\frac{m}{VE_{\mathbf{p}_1'}}} \sqrt{\frac{m}{VE_{\mathbf{p}_2'}}} \sqrt{\frac{m}{VE_{\mathbf{p}_2}}} \sqrt{\frac{m}{VE_{\mathbf{p}_1}}} (2\pi)^4\, \delta^{(4)}\left(p_1' + p_2' - \left(p_1 + p_2\right)\right) \cdot \\ \cdot \left\{\bar{w}_{r_2}\left(\mathbf{p}_2\right) \gamma_\mu w_{r_2'}\left(\mathbf{p}_2'\right) iD_F^{\mu\nu}\left(p_2' - p_2\right) \bar{u}_{r_1'}\left(\mathbf{p}_1'\right) \gamma_\nu u_{r_1}\left(\mathbf{p}_1\right)\right\} \tag{W.355}$$

---

[137] Remember that the propagator does not contain any deletion or creation operators.

[138] Remind the rules of matrix multiplication, $xMy = \sum x_i M_{ij} y_j$.

which we write as[139]

$$S_{B2}^{(2)} = \left(\prod_{\mathbf{p}}^{\text{ext. ferm.}} \sqrt{\frac{m}{VE_{\mathbf{p}}}} \right) (2\pi)^4 \, \delta^{(4)} \left(p_2 + p_1 - \left(p_2' + p_1' \right) \right) \mathcal{M}_{B2}^{(2)} \tag{W.356}$$

with

$$\mathcal{M}_{B2}^{(2)} = q^2 \bar{w}_{r_2} (\mathbf{p}_2) \, \gamma_\mu w_{r_2'} \left(\mathbf{p}_2' \right) i D_F^{\mu\nu} \left(p_2' - p_2 \right) \bar{u}_{r_1'} \left(\mathbf{p}_1' \right) \gamma_\nu u_{r_1} (\mathbf{p}_1) \, . \tag{W.357}$$

Again, the mathematical formulation (W.357) (from right to left) and the diagram (W.12) (in direction of the arrows) provide the same information: $u_{r_1} (\mathbf{p}_1)$ describes an incoming electron with quantum numbers $(r_1, \mathbf{p}_1)$, the vertex contributes $-iq\gamma_\nu$, the term $\bar{u}_{r_1'} \left(\mathbf{p}_1' \right)$ stands for an outgoing electron with $\left(r_1', \mathbf{p}_1' \right)$. A virtual photon with $k = p_2' - p_2$ is the link to $w_{r_2'} \left(\mathbf{p}_2' \right)$, creating a positron with $(r_2', \mathbf{p}_2')$, followed by the vertex $-iq\gamma_\mu$, and finally $\bar{w}_{r_2} (\mathbf{p}_2)$, annihilating a positron $(r_2, \mathbf{p}_2)$.

W.12.1.3 Total Bhabha Scattering

Now we can add up the two contributions for Bhaba scattering. Type 1 of (W.343) and (W.344), and type 2 of (W.356) and (W.357). The complete S-matrix element for the 2nd order ($n = 2$) Bhabha scattering is

$$S_{\text{Bhabha}}^{(2)} = S_{B1}^{(2)} + S_{B2}^{(2)} = \left(\prod_{\mathbf{p}}^{\text{ext. ferm.}} \sqrt{\frac{m}{VE_{\mathbf{p}}}} \right) (2\pi)^4 \, \delta^{(4)} \left(p_2 + p_1 - \left(p_2' + p_1' \right) \right) \mathcal{M}_{\text{Bhabha}}^{(2)} \tag{W.358}$$

with $\mathcal{M}_{\text{Bhabha}}^{(2)} = \mathcal{M}_{B1}^{(2)} + \mathcal{M}_{B2}^{(2)}$ and

$$\begin{aligned} \mathcal{M}_{B1}^{(2)} &= \left(-q^2 \right) \bar{u}_{r_1'} \left(\mathbf{p}_1' \right) \gamma_\mu w_{r_2'} \left(\mathbf{p}_2' \right) i D_F^{\mu\nu} \left(p_1 + p_2 \right) \bar{w}_{r_2} (\mathbf{p}_2) \, \gamma_\nu u_{r_1} (\mathbf{p}_1) \quad \text{; annihilation scattering} \\ \mathcal{M}_{B2}^{(2)} &= \left(q^2 \right) \bar{w}_{r_2} (\mathbf{p}_2) \, \gamma_\mu w_{r_2'} \left(\mathbf{p}_2' \right) i D_F^{\mu\nu} \left(p_2' - p_2 \right) \bar{u}_{r_1'} \left(\mathbf{p}_1' \right) \gamma_\nu u_{r_1} (\mathbf{p}_1) \quad \text{; direct scattering.} \end{aligned} \tag{W.359}$$

We see that the real work in QED is in the calculation of the Feynman amplitude - the mass or normalization factors as well as the conservation of energy and momentum (i.e., the delta function) are easily written down. Thus, it is a great relief that this centerpiece of QED (or QFT) can be done by means of the easy to read Feynman diagrams.

By the way: In general, only the squared Feynman amplitude $|\mathcal{M}|^2$ occurs in applications. It follows that e.g. the sign of q is irrelevant, but the relative sign between the subamplitudes matters, in this case $\mathcal{M}_{B1}^{(2)}$ and $\mathcal{M}_{B2}^{(2)}$.

[139] The formulation $\left(\prod_{\mathbf{p}}^{\text{ext. ferm.}} \sqrt{\frac{m}{VE_{\mathbf{p}}}} \right)$ instead of $\sqrt{\frac{m}{VE_{\mathbf{p}_1'}}} \sqrt{\frac{m}{VE_{\mathbf{p}_2'}}} \sqrt{\frac{m}{VE_{\mathbf{p}_2}}} \sqrt{\frac{m}{VE_{\mathbf{p}_1}}}$ allows also for muons, tauons and so on.

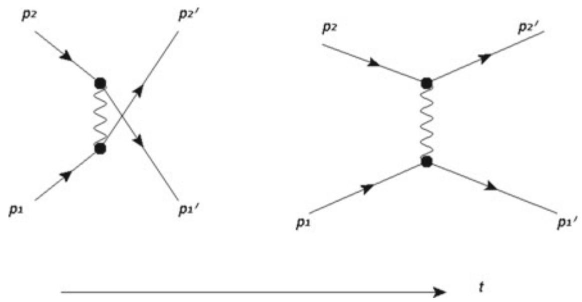


Fig. W.13 Moller scattering, type 1 (exchange scattering) and type 2 (direct scattering)

W.12.2 Møller Scattering, $e^-e^- \to e^-e^-$

Møller scattering is defined as electron-electron scattering. We will not run through the complete calculation as for Bhabha scattering. Again, we have two types of scattering which are characterized by the following terms (Fig. W.13):

$$\begin{array}{rl} \text{exchange scattering, type 1} & \left(\bar{\psi}^c_{1'}\gamma_\mu\psi^d_2\right)_1 \left(\bar{\psi}^c_{2'}\gamma_\nu\psi^d_1\right)_2 + \left(\bar{\psi}^c_{1'}\gamma_\mu\psi^d_2\right)_2 \left(\bar{\psi}^c_{2'}\gamma_\nu\psi^d_1\right)_1 \\ \text{direct scattering, type 2} & \left(\bar{\psi}^c_{2'}\gamma_\mu\psi^d_2\right)_1 \left(\bar{\psi}^c_{1'}\gamma_\nu\psi^d_1\right)_2 + \left(\bar{\psi}^c_{2'}\gamma_\mu\psi^d_2\right)_2 \left(\bar{\psi}^c_{1'}\gamma_\nu\psi^d_1\right)_1 \end{array}. \tag{W.360}$$

For type 1, we have a virtual photon with $k = p_1 - p_2'$ (or equivalently $k = p_2 - p_1'$), for type 2 a photon with $k = p_1 - p_1'$ (or $k = p_2 - p_2'$). For better identification, the field operators in (W.360) are indexed in the same manner as the momenta.

This time, we do not perform the calculations as we have done for Bhabha scattering, but transform the diagrams directly into the transition amplitude. In a first step, we formulate the conservation of the 4-momentum and the mass and volume factors. This gives

$$S^{(2)}_{\text{Møller}} = (2\pi)^4\, \delta^{(4)}\left(p_2 + p_1 - \left(p_2' + p_1'\right)\right) \left(\prod_{\mathbf{p}}^{\text{ext. ferm.}} \sqrt{\frac{m}{VE_{\mathbf{p}}}}\right) \mathcal{M}^{(2)}_{\text{Møller}} \tag{W.361}$$

with $\mathcal{M}^{(2)}_{\text{Møller}} = \mathcal{M}^{(2)}_{M1} + \mathcal{M}^{(2)}_{M2}$ where $\mathcal{M}^{(2)}_{M1}$ and $\mathcal{M}^{(2)}_{M2}$ are the Feynman amplitudes for the two processes. As explained above, we have to add the amplitudes, not their squares, because we can not know which one of the two possibilities may have occurred.

The amplitudes themselves are read off from the diagrams. We repeat that for comparing the formulas with the diagrams, one has to read the formulas from right to left and the diagrams in direction of the arrows. For type 1, an incoming electron with $(\mathbf{p}_2, r_2)$ is switched into an outgoing electron with $\left(\mathbf{p}_1', r_1'\right)$ and the other one with $(\mathbf{p}_1, r_1)$ into $\left(\mathbf{p}_2', r_2'\right)$; for the virtual photon we have $k = p_1 - p_2'$. For type 2, there is an analogous formulation. Incorporating the vertices by $-iq\gamma_\mu$, we arrive at the Feynman amplitudes for Møller scattering:

$$\begin{aligned} \mathcal{M}_{M1}^{(2)} &= \left(q^2\right) \bar{u}_{r_2'}\left(\mathbf{p}_2'\right) \gamma_\mu u_{r_1}\left(\mathbf{p}_1\right) iD_F^{\mu\nu}\left(p_2 - p_1'\right) \bar{u}_{r_1'}\left(\mathbf{p}_1'\right) \gamma_\nu u_{r_2}\left(\mathbf{p}_2\right) \quad \text{; exchange scattering} \\ \mathcal{M}_{M2}^{(2)} &= \left(-q^2\right) \bar{u}_{r_1'}\left(\mathbf{p}_1'\right) \gamma_\mu u_{r_1}\left(\mathbf{p}_1\right) iD_F^{\mu\nu}\left(p_2 - p_2'\right) \bar{u}_{r_2'}\left(\mathbf{p}_2'\right) \gamma_\nu u_{r_2}\left(\mathbf{p}_2\right) \quad \text{; direct scattering.} \end{aligned} \tag{W.362}$$

W.12.3 Scattering Cross Section and Feynman Amplitude

In this way the Feynman amplitudes can be read directly from the corresponding diagrams. In principle, this holds for all processes in Figs. W.7, W.8, W.9, W.10 and W.11. However, for the diagrams with loops there is in addition a special feature which we will discuss only in the last section.

What remains is the explicit evaluation of the Feynman amplitudes as given e.g. for Bhabha and for Møller scattering by (W.359) and (W.362). To this end we have to insert and use the definitions and properties of the spinors $u_r(\mathbf{p})$ and $w_r(\mathbf{p})$ as given in (W.94)–(W.99) in section 'Quantization of free fields - Dirac'. The evaluation looks like an quite innocent task, but actually it is a very tedious and lengthy affair[140] which would consume too much space here. Thus, we omit it and refer to the literature. But we give now the results for Bhabha and Møller scattering in the form of the corresponding scattering cross sections.

W.12.3.1 Scattering of Two Particles

In scattering experiments, the (differential) scattering cross section $\frac{d\sigma}{d\Omega}$(cf. Vol. 2, Chap. 25) is central. We now look for the relation between $\frac{d\sigma}{d\Omega}$ and the Feynman amplitude $\mathcal{M}$.

We assume that there are two initial and two final particles. Then one can show, that in the center of mass system holds[141]

[140]The calculation is based to a large extent on trace techniques.

[141]A remark concerning the unit of the scattering cross section which should have units of area, i.e., length squared. In natural units ($\hbar = 1, c = 1$; see Vol. 1 Appendix B 'Units and Constants'), energy and mass have the same physical unit, since $E = mc^2 \rightarrow E = m$. The same holds for frequency ω and wave number k. Thus, $[E] = [m] = [\omega] = [k]$. Due to $\lambda = h/p \rightarrow \lambda = 1/p$ follows $[\lambda] =$ length $= 1/[E]$ or $[E] = 1/$length. Since $D_F^{\mu\nu} \sim 1/k^2$ we have $[\mathcal{M}] = \left[D_F^{\mu\nu}\right] = 1/[k]^2 = 1/[E]^2$. It follows from (W.364)

$$\left[\frac{d\sigma}{d\Omega}\right] = \left[\frac{1}{E}\right]^2 \left[\frac{|\mathbf{p}_1'|}{|\mathbf{p}_1|}\right] [m]^4 \cdot \left[|\mathcal{M}|^2\right] = [E]^2 [\mathcal{M}]^2 = [E]^2 \frac{1}{[E]^4} = \frac{1}{[E]^2} = \text{length}^2 = \text{area} \tag{W.363}$$

as it should be.

$$\frac{d\sigma}{d\Omega} = \frac{1}{4} \frac{1}{(4\pi)^2 (E_1 + E_2)^2} \frac{|\mathbf{p}'_1|}{|\mathbf{p}_1|} \prod_i^{\text{ext. fermions}} (2m_i) \cdot |\mathcal{M}|^2 . \qquad \text{(W.364)}$$

The 4-momenta of the two particles before and after the scattering are given p_i and p'_i. Note that in the center of mass system holds $\mathbf{p}_2 = -\mathbf{p}_1$. Equation (W.364) is valid for arbitrary fermions (electrons, muons, tauons and their antiparticles) in the incoming and outgoing channel, e.g. for the process $e^+ + e^- = \mu^+ + \mu^-$.

The essential point is the direct proportionality of $\frac{d\sigma}{d\Omega}$ and $|\mathcal{M}|^2$. In this respect, the Feynman amplitude resembles the scattering amplitude.[142] For the considered examples, Bhabha and Møller scattering, we have in each case two different processes with different subamplitudes and thus $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2$. This means that we have a squared amplitude for each subprocess plus an additional interference term,

$$|\mathcal{M}|^2 = |\mathcal{M}_1|^2 + |\mathcal{M}_2|^2 + 2\,\mathrm{Re}\,\mathcal{M}_1\mathcal{M}_2^*. \qquad \text{(W.365)}$$

All following results are calculated assuming spin-averaging. This means that (1) that the spin orientations of the incoming particles are random and (2) the spin directions of the outgoing particles are not measured.[143] Thus, the cross section is a mean value of the cross sections (W.364) for all possible spin directions, i.e.,

$$|\mathcal{M}|^2 \to \frac{1}{4} \sum_{r_1, r_2, r'_1, r'_2} |\mathcal{M}|^2 . \qquad \text{(W.366)}$$

As said above, the evaluation of the Feynman amplitudes would unfortunately consume too much place here. Thus, we report in the following just the results. The graphical representation is found in Fig. W.14.

W.12.3.2 Møller Scattering

In case of Møller scattering, the general expression for the (unpolarized) scattering cross section reads[144]:

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{Møller}} = \frac{\alpha^2}{16E^2\mathbf{p}^4} \cdot \left\{ \begin{array}{c} \frac{1}{\sin^4\frac{\vartheta}{2}} \left[m^4 + 4\mathbf{p}^2 m^2 \cos^2\frac{\vartheta}{2} + 2\mathbf{p}^4\left(1 + \cos^4\frac{\vartheta}{2}\right)\right] + \\ + \frac{1}{\cos^4\frac{\vartheta}{2}} \left[m^4 + 4\mathbf{p}^2 m^2 \sin^2\frac{\vartheta}{2} + 2\mathbf{p}^4\left(1 + \sin^4\frac{\vartheta}{2}\right)\right] + \\ + \frac{1}{\cos^2\frac{\vartheta}{2}\sin^2\frac{\vartheta}{2}}\left(4\mathbf{p}^4 - m^4\right) \end{array} \right\} \qquad \text{(W.367)}$$

142Also decay rates for decay processes (e.g. $K^0 \to \pi^+\pi^-$) are proportional to $|\mathcal{M}|^2$.

143Note that a particle beam with random spin orientations is often called ‘unpolarized’. Thus, the term ‘polarization’ is used to include both fermion spin and photon polarizations.

144α is the fine structure constant, $\alpha = q^2/4\pi$.

where ϑ is the scattering angle. One by one, the summands are contributions of (1) direct scattering (type 2), (2) exchange scattering (type 1) and (3) of the interference between the two scattering processes.

One can rewrite the scattering cross section to give

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{Møller}} = \frac{\alpha^2\left(2E^2-m^2\right)^2}{4E^2\left(E^2-m^2\right)^2}\left[\frac{4}{\sin^4\vartheta}-\frac{3}{\sin^2\vartheta}+\frac{\left(E^2-m^2\right)^2}{\left(2E^2-m^2\right)^2}\left(1+\frac{4}{\sin^2\vartheta}\right)\right]. \tag{W.368}$$

In the high relativistic limit $\mathbf{p}^2 \gg m^2$, we have

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{Møller},\ \mathbf{p}^2\gg m^2} = \frac{\alpha^2}{8E^2}\left\{\frac{1+\cos^4\frac{\vartheta}{2}}{\sin^4\frac{\vartheta}{2}}+\frac{1+\sin^4\frac{\vartheta}{2}}{\cos^4\frac{\vartheta}{2}}+\frac{2}{\cos^2\frac{\vartheta}{2}\sin^2\frac{\vartheta}{2}}\right\} = \frac{\alpha^2}{4E^2}\frac{\left(3+\cos^2\vartheta\right)^2}{\sin^4\vartheta} \tag{W.369}$$

and in the nonrelativistic limit $\mathbf{p}^2 \ll m^2$ follows

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{Møller},\ \mathbf{p}^2\ll m^2} = \frac{\alpha^2 m^2}{16\mathbf{p}^4}\left[\frac{1}{\sin^4\frac{\vartheta}{2}}+\frac{1}{\cos^4\frac{\vartheta}{2}}-\frac{1}{\sin^2\frac{\vartheta}{2}\cos^2\frac{\vartheta}{2}}\right] = \frac{\alpha^2 m^2}{4\mathbf{p}^4}\left[\frac{1+3\cos^2\vartheta}{\sin^4\vartheta}\right]. \tag{W.370}$$

Comparing this with the classical (i.e., nonrelativistic) Rutherford cross section (scattering of identical particles)

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{Rutherford}} = \frac{\alpha^2 m^2}{16\mathbf{p}^4}\left[\frac{1}{\sin^4\frac{\vartheta}{2}}+\frac{1}{\cos^4\frac{\vartheta}{2}}\right] \tag{W.371}$$

shows that QED gives an additional interference term. (See also Appendix O, Vol. 2, ‘Scattering of identical particles’).

W.12.3.3 Bhabha Scattering

In case of Bhabha scattering, the general expression for the scattering cross section reads

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{Bhabha}} = \frac{\alpha^2}{16E^2\mathbf{p}^4}\cdot\left\{\begin{array}{c}\frac{1}{\sin^4\frac{\vartheta}{2}}\left[m^4+4\mathbf{p}^2m^2\cos^2\frac{\vartheta}{2}+2\mathbf{p}^4\left(1+\cos^4\frac{\vartheta}{2}\right)\right]+\\ +\frac{\mathbf{p}^4}{E^4}\left[3m^4+4\mathbf{p}^2m^2+2\mathbf{p}^4\cos^2\frac{\vartheta}{2}\right]-\\ -\frac{\mathbf{p}^2}{E^2\sin^2\frac{\vartheta}{2}}\left[3m^4+8\mathbf{p}^2m^2\cos^2\frac{\vartheta}{2}+4\mathbf{p}^4\cos^4\frac{\vartheta}{2}\right]\end{array}\right\}. \tag{W.372}$$

One by one, the summands are contributions of (1) direct scattering (type 2), (2) annihilation scattering (type 1) and (3) of the interference between the two scattering processes.

In the high relativistic limit with $\mathbf{p}^2 \gg m^2$, we have

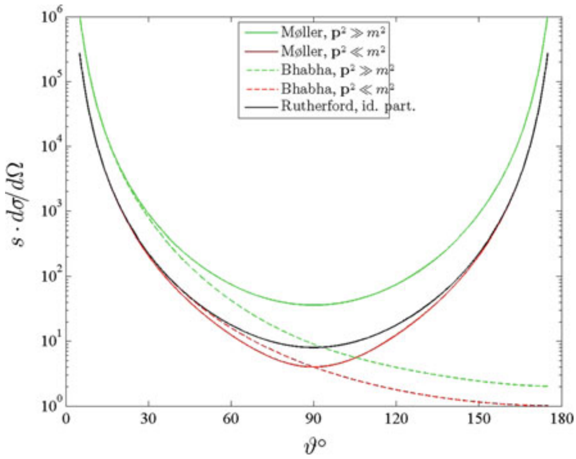


Fig. W.14 For better comparability, the cross sections are multiplied by $s = \frac{16\mathbf{p}^4}{\alpha^2 E^2}$

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right|_{\text{Bhabha, }\mathbf{p}^2 \gg m^2} = \frac{\alpha^2}{8E^2}\left\{\frac{1+\cos^4\frac{\vartheta}{2}}{\sin^4\frac{\vartheta}{2}} + \cos^2\frac{\vartheta}{2} - \frac{2\cos^4\frac{\vartheta}{2}}{\sin^2\frac{\vartheta}{2}}\right\} = \frac{\alpha^2}{8E^2}\frac{\left(3+\cos^2\vartheta\right)^2}{2\left(1-\cos\vartheta\right)^2}. \tag{W.373}$$

In the nonrelativistic limit $\mathbf{p}^2 \ll m^2$ follows

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right|_{\text{Bhabha, }\mathbf{p}^2 \ll m^2} = \frac{\alpha^2 m^2}{16\mathbf{p}^4}\frac{1}{\sin^4\frac{\vartheta}{2}}. \tag{W.374}$$

The contribution of the annihilation process and the interference is suppressed, since at low energies they are of order $O(\frac{\mathbf{p}^2}{m^2})$. As is seen, the cross section agrees with the classical Rutherford cross section (scattering of non-identical particles):

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right|_{\text{Rutherford}} = \frac{\alpha^2 m^2}{16\mathbf{p}^4}\frac{1}{\sin^4\frac{\vartheta}{2}} \tag{W.375}$$

W.12.4 Exercises and Solutions

1. Prove (W.340).
   Solution: We start with

$$X = \int\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2\, D_F^{\mu\nu}\left(x_1 - x_2\right) e^{iP_1' x_1} e^{-iPx_2}\ ;\ P = p_1 + p_2\ ;\ P' = p_1' + p_2'. \tag{W.376}$$

   Inserting the propagator ((W.286) in sec. 'contraction, propagator, Wick's theorem') brings

$$\begin{aligned} X &= -\int\int \mathrm{d}^4x_1\, \mathrm{d}^4x_2\; g^{\mu\nu} \tfrac{1}{(2\pi)^4} \int \mathrm{d}^4k\; \tfrac{1}{k^2+i\varepsilon} e^{-ik(x_1-x_2)} e^{iP_1'x_1} e^{-iPx_2} = \\ &= -g^{\mu\nu} \tfrac{1}{(2\pi)^4} \int \mathrm{d}^4x_1\; \int \mathrm{d}^4k\; \tfrac{1}{k^2+i\varepsilon} e^{-ikx_1} e^{iP_1'x_1} \int \mathrm{d}^4x_2\; e^{ikx_2} e^{-iPx_2} = \\ &= -g^{\mu\nu} \tfrac{1}{(2\pi)^4} \int \mathrm{d}^4x_1\; \int \mathrm{d}^4k\; \tfrac{1}{k^2+i\varepsilon} e^{-ikx_1} e^{iP_1'x_1}\, (2\pi)^4\, \delta^{(4)}\,(k-P) = \\ &= -g^{\mu\nu} \int \mathrm{d}^4x_1\; \tfrac{1}{(p_1+p_2)^2+i\varepsilon} e^{-iPx_1} e^{iP_1'x_1} = -g^{\mu\nu}\,(2\pi)^4\, \tfrac{\delta^{(4)}\left(P-P'\right)}{P^2+i\varepsilon} = \\ &= -g^{\mu\nu} \tfrac{1}{P^2+i\varepsilon}\,(2\pi)^4\, \delta^{(4)}\left(P-P'\right) = D_F^{\mu\nu}\,(P)\,(2\pi)^4\, \delta^{(4)}\left(P-P'\right) \end{aligned} \tag{W.377}$$

where $D_F^{\mu\nu}\,(p_1+p_2)$ is the Fourier transform of the propagator, i.e.,

$$D_F^{\mu\nu}\,(P) = -g^{\mu\nu} \frac{1}{P^2+i\varepsilon}. \tag{W.378}$$

2. Formulate the Feynman amplitudes and the S-matrix for the processes in Fig. W.8.

Solution: (1) Compton scattering. Labels are: incoming electron p, incoming photon k, outgoing electron p', outgoing photon k'. It follows

$$\mathcal{M}_{C1}^{(2)} = -q^2 \varepsilon_{\mu,r'}\left(\mathbf{k}'\right) \bar{u}_{s'}\left(\mathbf{p}'\right) \gamma^\mu i S_F\,(q=p+k)\,\varepsilon_{\nu,r}\,(\mathbf{k})\,\gamma^\nu u_s\,(\mathbf{p}) \tag{W.379}$$

and

$$\mathcal{M}_{C2}^{(2)} = -q^2 \varepsilon_{\mu,r}\,(\mathbf{k})\, \bar{u}_{s'}\left(\mathbf{p}'\right) \gamma^\mu i S_F\left(q=p-k'\right) \varepsilon_{\nu,r'}\left(\mathbf{k}'\right) \gamma^\nu u_s\,(\mathbf{p}) \tag{W.380}$$

and

$$S_C^{(2)} = \sqrt{\frac{m}{VE_{\mathbf{p}}}}\sqrt{\frac{m}{VE_{\mathbf{p}'}}}\sqrt{\frac{1}{2VE_{\mathbf{k}}}}\sqrt{\frac{1}{2VE_{\mathbf{k}'}}}\,(2\pi)^4\,\delta\left(p+k-p'-k'\right)\left(\mathcal{M}_{C1}^{(2)}+\mathcal{M}_{C2}^{(2)}\right). \tag{W.381}$$

(2) Pair annihilation. Labels are: electron p_1, positron p_2. Photon at the electron-electron or positron-positron vertex k_1, photon at the electron-positron vertex k_2. It follows

$$\mathcal{M}_{PA1}^{(2)} = -q^2 \varepsilon_{\mu,r_2}\,(\mathbf{k}_2)\,\varepsilon_{\nu,r_1}\,(\mathbf{k}_1)\,\bar{w}_{r_2}\,(\mathbf{p}_2)\,\gamma^\mu i S_F\,(p_1-k_1)\,\gamma^\nu u_{s_1}\,(\mathbf{p}_1) \tag{W.382}$$

and

$$\mathcal{M}_{PA2}^{(2)} = -q^2 \varepsilon_{\mu,r_1}\,(\mathbf{k}_1)\,\varepsilon_{\nu,r_2}\,(\mathbf{k}_2)\,\bar{w}_{r_2}\,(\mathbf{p}_2)\,\gamma^\mu i S_F\,(p_1-k_2)\,\gamma^\nu u_{s_1}\,(\mathbf{p}_1) \tag{W.383}$$

and

$$S_{PA}^{(2)} = \sqrt{\frac{m}{VE_{\mathbf{p}_1}}}\sqrt{\frac{m}{VE_{\mathbf{p}_2'}}}\sqrt{\frac{1}{2VE_{\mathbf{k}_1}}}\sqrt{\frac{1}{2VE_{\mathbf{k}_2'}}}\,(2\pi)^4\,\delta\,(p_1+p_2-k_1-k_2)\left(\mathcal{M}_{PA1}^{(2)}+\mathcal{M}_{PA2}^{(2)}\right). \tag{W.384}$$

3. Show

$$\frac{\alpha^2 m^2}{4\mathbf{p}^4}\frac{1+3\cos^2\vartheta}{\sin^4\vartheta}=\frac{\alpha^2 m^2}{16\mathbf{p}^4}\left[\frac{1}{\sin^4\frac{\vartheta}{2}}+\frac{1}{\cos^4\frac{\vartheta}{2}}-\frac{1}{\sin^2\frac{\vartheta}{2}\cos^2\frac{\vartheta}{2}}\right]. \tag{W.385}$$

Solution:

$$\begin{aligned}\text{r.h.s.} &= \frac{\alpha^2 m^2}{16\mathbf{p}^4}\left[\frac{1}{\sin^4\frac{\vartheta}{2}}+\frac{1}{\cos^4\frac{\vartheta}{2}}-\frac{1}{\sin^2\frac{\vartheta}{2}\cos^2\frac{\vartheta}{2}}\right]=\\ &=\frac{\alpha^2 m^2}{16\mathbf{p}^4}\frac{\cos^4\frac{\vartheta}{2}+\sin^4\frac{\vartheta}{2}-\sin^2\frac{\vartheta}{2}\cos^2\frac{\vartheta}{2}}{\sin^4\frac{\vartheta}{2}\cos^4\frac{\vartheta}{2}}.\end{aligned} \tag{W.386}$$

With $\sin\frac{\vartheta}{2}=\sqrt{\frac{1}{2}\left(1-\cos\vartheta\right)}$ and $\cos\frac{\vartheta}{2}=\sqrt{\frac{1}{2}\left(1+\cos\vartheta\right)}$ follows

$$\begin{aligned}\text{r.h.s.} &= \frac{\alpha^2 m^2}{16\mathbf{p}^4}\left[\frac{\left(\frac{1}{2}\right)^2(1+\cos\vartheta)^2+\left(\frac{1}{2}\right)^2(1-\cos\vartheta)^2-\left(\frac{1}{2}\right)^2\left(1-\cos^2\vartheta\right)}{\left(\frac{1}{2}\right)^4\left(1-\cos^2\vartheta\right)^2}\right]=\\ &=\frac{\alpha^2 m^2}{16\mathbf{p}^4}4\frac{1+3\cos^2\vartheta}{\sin^4\vartheta}=\text{l.h.s.}\end{aligned} \tag{W.387}$$

W.13 High Precision and Infinities

As mentioned in the introduction to this chapter, we tried to set some stepping stones on the way to the realm of QFT, especially QED. Of course, we had to leave questions open, and there are some gaps. Nevertheless, we have accumulated a lot of material. As it turned out, the triad Feynman amplitudes, diagrams and rules is a centerpiece of QED.

In this last section, we want to complete the Feynman rules and show how (in principle) to formulate S-matrix elements of arbitrary order. After that, we discuss why QED is considered the most stringently proven theory in physics. And finally, we cursorily consider renormalization, i.e., how to handle divergent integrals which emerge in the calculation of so-called loop diagrams.

W.13.1 Feynman Rules, Diagrams, Amplitudes for QED

We have seen in the previous sections that there is a one-to-one relation between Feynman diagrams and Feynman amplitudes.[145] The relation is given by the Feynman rules most of which we have already formulated. However, three are still lacking and are presented now. We will not derive them, since our main intention here is to

[145] Feynman diagrams are also used in other fields as e.g. quantum chromodynamics (QCD). In principle it is the same technique though there are new elements as e.g. gluons and ghosts and there is another coupling constant; instead of the fine structure constant $\alpha_{QED}\approx 1/137$ we have $\alpha_{QCD}\sim 1$.

complete the whole set of rules. However, below are found examples that at least illustrate how the new rules are used.

We can classify the diagrams by the order n of the scattering matrix $S^{(n)}$ which equals the number of vertices. The rules formulated below are valid for all n.

Assume we have a Feynman diagram of order n. Then we can determine the Feynman amplitude of this diagram according to the following Feynman rules for QED.

1. For each vertex, there is a factor $-iq\gamma^{\mu}$ with $q = \sqrt{4\pi\alpha}$.
2. For each external line, there is one of the following spinors and polarization vectors:
   (a) for each incoming electron $u_r(\mathbf{p})$, for each outgoing electron $\bar{u}_r(\mathbf{p})$,
   (b) for each incoming positron $\bar{w}_r(\mathbf{p})$, for each outgoing positron $w_r(\mathbf{p})$,
   (c) for each incoming photon $\varepsilon_{\lambda\mu}(\mathbf{k})$, for each outgoing photon $\varepsilon_{\lambda\mu}(\mathbf{k})$.
3. For each internal
   (a) photon line (photon propagator), there is a factor $iD_F^{\mu\nu}(k) = i\frac{-g^{\mu\nu}}{k^2+i\varepsilon}$.
   (b) fermionic line (fermion propagator), there is a factor $iS_F(p) = i\frac{\not{p}+m}{p^2-m^2+i\varepsilon}$.
4. At each vertex, the 4-momenta are conserved.
5. The spinor factors (γ matrices, fermionic propagators, 4-spinors) are ordered as follows: reading the formulas from right to left, they occur in the same sequence as following the fermion line in the direction of its arrows through the vertex. Note that order is crucial as it affects the spinor matrix multiplication.
6. For each interchange of neighboring fermionic operators which is required to bring the expression in appropriate normal order, one has to multiply the expression by a factor of (-1). The word 'appropriate' means the following: if there are several subamplitudes (as in case of Bhabha scattering), each subamplitude must be in the same normal order.
7. For the interaction of a charged particle with a static external field, the vector potential is given by $A_{\mu}^{e}(\mathbf{k}) = \int d^3x\, e^{-i\mathbf{kx}} A_{\mu}^{e}(\mathbf{x})$. In addition, the delta function $(2\pi)^4\,\delta^{(4)}\left(\sum p_f - \sum p_i\right)$ in S_{fi} is replaced by $2\pi\delta\left(\sum E_f - \sum E_i\right)$.[146]

   We have derived these rules in the preceding sections. As mentioned above, there are three more rules which have to do with loops and renormalization, i.e., with redefined values of mass and charge, see below. These rules are:
8. For each closed loop of fermion lines, one has to take the trace in spinor space of the resulting matrix and multiply by a factor of (-1).
9. For each 4-momentum p which is not fixed by 4-momentum conservation, one has to carry out the integration $\frac{1}{(2\pi)^4}\int d^4p$.[147]

[146] In other words: for external fields, the rules of the game are changed to some extent.

[147] Rule 8 and 9 are illustrated below by means of Bhabha scattering plus photon loop.

10. Renormalization of the mass leads to a mass counterterm diagram and, in the amplitudes, to an additional term equal to $i\delta m$.[148]

It is a very compact set - two handfuls of rules connect biuniquely the graphical and mathematical representation of all physical processes of QED.

Following these rules we can write down the exact analytical expression for the Feynman amplitude from the diagram.[149] In general, for a given order n there are several subamplitudes $\mathcal{M}_i^{(n)}$ as e.g. in the case of Bhabha scattering where we found two different processes contributing to $\mathcal{M}_{\text{Bhabha}}^{(2)}$. Via $\mathcal{M}^{(n)} = \sum_i \mathcal{M}_i^{(n)}$ they contribute to the Feynman amplitude $\mathcal{M}^{(n)}$.

The next question is if we want to examine the scattering of order n - then only $\mathcal{M} = \mathcal{M}^{(n)}$ is considered - or the entire process up to and including the order N - then $\mathcal{M} = \sum_{n=1}^{N} \mathcal{M}^{(n)}$ is taken into account. Finally, the S-matrix element $\langle f|\, S\,|i\rangle = S_{fi}$ is given by (discrete version)

$$S_{fi} = \delta_{fi} + \left[(2\pi)^4\, \delta^{(4)}\left(P_f - P_i\right) \cdot \prod^{\text{ext. phot.}} \sqrt{\frac{1}{2VE_{\mathbf{k}}}} \cdot \prod^{\text{ext. ferm.}} \sqrt{\frac{m}{VE_{\mathbf{p}}}} \right] \cdot \mathcal{M}. \quad \text{(W.388)}$$

As mentioned above, we will not explicitly calculate the Feynman amplitudes by inserting and evaluating the spinors u and w and the polarization vectors ε. This simply would be too lengthy and can not be afforded within the scope of this brief introduction to QFT.

W.13.2 Extraordinary Precision

The high-precision calculation of the g-factor (Landé factor) is one of the great success stories of QED. The Dirac equation predicts the magnetic moment of the electron $\boldsymbol{\mu}_e$ to be given by (see Appendix U, Vol. 1)

$$\boldsymbol{\mu}_e = g \frac{q}{2m} \mathbf{s} \ ; \ g = 2. \quad \text{(W.389)}$$

QED allows for an impressively accurate determination of the anomalous magnetic moment, i.e., the deviation from $g = 2$. In 1948, Julian Schwinger[150] showed that radiation and re-absorption of a single virtual photon (see e.g. Fig. W.15 for an electron in an external magnetic field) modifies the g-value from 2 to $2\left(1 + \frac{\alpha}{2\pi}\right) \approx 2,00332$. Incorporating effects of order α^3 leads to $g = 2,002\,319\,304$. Meanwhile (April 2018), taking into account more and more higher order corrections (up to

[148] For δm and a short explanation of the rule see below.

[149] Of course one can still calculate the amplitudes by the purely formal way, i.e., by inserting the field operators into $S^{(2)}$. But that is much more complicated and lengthy, as we have seen.

[150] Schwinger, Julian Seymour, 1918–1994, US-American physicist, nobel prize 1965.

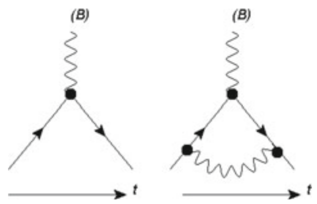


Fig. W.15 Vertex correction (one loop contribution) of the magnetic moment; electron in an e.g. external magnetic field

α^5), the theoretical value is given by 2, 002 319 304 363 286(1528), whereby the experimental value is 2, 002 319 304 361 46(56).[151] The current theoretical limit is due to the fact that for higher orders n there is a rapidly increasing number of integrals to evaluate (3th order 72 diagrams, 4th order 981, 5th order 12672).

As is seen, the theoretical and experimental values agree to more than 10 significant figures. This makes the magnetic moment of the electron the most precisely determined physical quantity and, hence, QED the most thoroughly tested theory.[152]

W.13.3 *Problematic Loops, Infinities, Renormalization*

The extraordinary precision is considered proof by (almost) all physicists that QED is correct, although there are significant problems with infinities which we will present now. The origin are three diagrams containing loops which lead to divergent integrals. These infinities may be accounted for by a redefinition of mass and charge, called *renormalization*. This is a central issue of QED - in any textbook of QED, a substantial part is dedicated to this (quite technical) topic. Within the framework of our compressed considerations we can only cursorily touch on the topic.

W.13.3.1 Three Problematic Loops

In QED, there are three types of problematic diagrams, illustrated in Fig. W.16. We have (1) ‘bubble propagators’: a real or virtual photon creates a fermion/antifermion pair (also called photon loop, photon self energy, vacuum polarization, closed fermion loop, (2) ‘dressed fermions’: a real or virtual fermion emits and reabsorbs a virtual photon (fermion loop, fermion self energy) and (3) vertex correction: a virtual photon connects fermions across a previous vertex (vertex loop correction, vertex modifi-

[151]From https://en.wikipedia.org/wiki/Anomalous_magnetic_dipole_moment. The National Institute of Standards and Technology gives for the experimental value 2, 002 319 304 361 82(52). Current status april 2018.

[152]The myon has a higher mass than the electron ($m_\mu \approx 200 m_e$). Its anomalous magnetic moment is therefore sensitive to interactions from physics beyond QED, even beyond the Standard Model. Thus, it may be used as a probe for new physics.

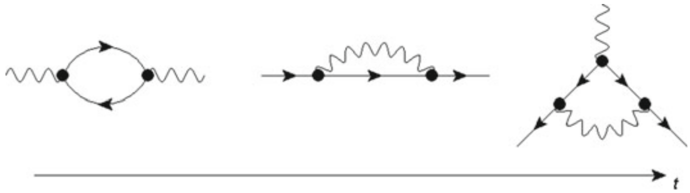


Fig. W.16 Photon self energy, fermion self energy and vertex loop correction

cation). These three diagrams lead to divergent integrals.[153] The reason is that the momentum p in the loops is not fixed which implies integration over p, see Feynman rule 9. However, as it turns out, these integrals are divergent. Such integrals are pretty unwelcome in any theory and usually imply the death of that theory, but not in QED.

W.13.3.2 Renormalization, Basic Idea

In fact, these problems are healed by a redefinition of mass and charge, i.e., by renormalization. The redefinition is chosen in such a way that it compensates for the infinities and gives us finite answers. Following this idea leads to two kinds of charges and masses. There are the bare charge and bare mass q_0 and m_0. We cannot measure them since nature shows us only the 'ordinary' measurable quantities q and m (or q_{renorm} and m_{renorm}), i.e., the renormalized quantities. They are so to speak real-world versions of the bare items, dressed up with corrections, i.e., contributions from one or several loops (Fig. W.17).

W.13.3.3 Renormalization, Example

To illustrate the situation, we consider Bhabha scattering type 1 (annihilation scattering) plus a photon loop, see Fig. W.17. Following the Feynman rules, we can write down the transition amplitude for this scattering as (see rules 8 and 9):

$$\begin{aligned}\mathcal{M}_{B1}^{\text{photon loop}} &= -\tfrac{q^4}{(2\pi)^4}\bar{u}_{r_1'}\left(\mathbf{p}_1'\right)\gamma^\rho w_{r_2'}\left(\mathbf{p}_2'\right)D_{F\rho\eta}(k)\,\cdot \\ &\cdot\left\{Tr\int S_F(p)\,\gamma^\mu S_F(p-k)\,\gamma^\eta \mathrm{d}^4p\right\}D_{F\mu\nu}(k)\,\bar{w}_{r_2}(\mathbf{p}_2)\,\gamma^\nu u_{r_1}(\mathbf{p}_1)\,.\end{aligned} \tag{W.390}$$

It is the integral which causes problems:

$$I^{\mu\eta} = \int S_F(p)\,\gamma^\mu S_F(p-k)\,\gamma^\eta \mathrm{d}^4p = \int \frac{\not{p}+m}{p^2-m^2+i\varepsilon}\gamma^\mu\frac{\not{p}-\not{k}+m}{(p-k)^2-m^2+i\varepsilon}\gamma^\eta \mathrm{d}^4p. \tag{W.391}$$

[153]Remember that we had to introduce normal ordering due to infinite vacuum energies. It seems that QED is indeed plagued by infinities.

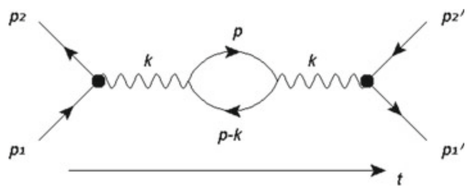


Fig. W.17 Bhabha scattering plus photon loop

Since the evaluation is rather lengthy; we just state the result: this integral *diverges* with $\ln p$ for high p.

Inspection of (W.390) shows that one can compensate for this infinity by e.g. (1) introducing a cut-off of the 4-momentum (the cut-off is often labeled Λ or Π, the method is called 'regularization')[154] and (2) redefining the charge in such a way that it eliminates the problematic terms. The bottom line is that renormalizing the three loop integrals leads e.g. to a renormalized charge q or q_{renorm} which is related to the bare charge q_0 by[155]

$$q = q_0 \left[1 - \frac{\alpha}{3\pi} \ln \left(\frac{\Lambda}{m} \right)^2 + \mathcal{O}(\alpha^2) \right]. \tag{W.392}$$

Note that the bare charge q_0 is greater than the renormalized, i.e., measurable charge q_{renorm}.[156]

Analogously, also the mass is renormalized which leads to changes in the Feynman amplitude and diagram. The reason is that the Lagrangian is formulated with the bare mass m_0. Substituting $m_0 = m - \delta m$ leads to an additional term $\delta m \bar{\psi} \psi$ (the mass counterterm of the Lagrangian; no photon involved, but sort of fermion self interaction) subsumed in the interaction part of the Lagrangian $\mathcal{L}^{\text{interaction}}$ and, hence, of the Hamiltonian $\mathcal{H}^{\text{interaction}} = -\mathcal{L}^{\text{interaction}}$. Thus, we have an extra term $\sim i\delta m$ in the amplitude and an extra diagram with an incoming and an outgoing fermion, usually represented by a fermionic line marked by a little centered cross '$\times$'.

W.13.3.4 Renormalization, Higher Orders and Question

Obviously, there are also loops in higher order terms. Figure W.18 shows as example the three (of infinite many ones) corrections for the photon propagator. One can show that renormalization can be performed in all orders and that the sum of higher order QED corrections converges.

[154]Remember that we have discussed a cut-off also in the context of infinite vacuum energies in the section 'Operator ordering'.

[155]For QED it is usually sufficient to ignore terms $\mathcal{O}(\alpha^2)$.

[156]As a consequence of renormalization, the value of the coupling 'constant' α (remind $q^2 = 4\pi\alpha$) depends on energy. As for numerical values: α increases with energy from $\alpha = 1/137$ at low energies to $\alpha = 1/128$ at 100 GeV. Therefore α is called *running coupling constant*.

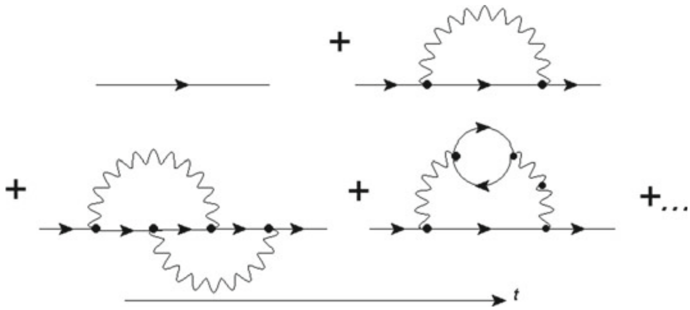


Fig. W.18 Fermion propagator with higher order divergent corrections

W.13.3.5 Renormalization, Vividly

We can get a vivid picture by looking at Fig. W.18. As is seen, the electron (or, more generally, fermion, real or virtual) constantly emits and reabsorbs photons. Hence, it is surrounded by a cloud of virtual photons (it is 'dressed') which carry a certain portion of energy and thus mass. Photons themselves (real and virtual ones) propagate while tearing fermion pairs from the vacuum. Thus, they are dressed by a cloud of fermion/antifermion pairs which modify the photon's amplitude for propagation between two points.

So we have fermions which emit and reabsorb photons which tear fermion pairs which emit and reabsorb photons which tear fermion pairs ... and so on, sort of a matryoshka doll.

We see that the vacuum is not a 'nothing', but an interacting medium. Dressing the bare photon in a cloud of fermion/antifermion pairs may be compared with a vacuum which interacts as a dielectric. In this context, one speaks of vacuum polarization or screening effect. We note that this effect can be and has been measured (affects the Lamb shift in the Hydrogen spectrum).

W.13.4 Conclusion

For many, QED is just a good and high-precision theory; admittedly, there occur problems, but they are manageable. Others feel uncomfortable and are rather unhappy about this juggling with infinities So were, among others also the founding fathers Dirac and Feynman. Dirac criticized: "This so-called 'good theory' does involve neglecting infinities which appear in its equations, neglecting them in an arbitrary way. This is just not sensible mathematics. Sensible mathematics involves neglecting a quantity when it is small – not neglecting it just because it is infinitely great and you do not want it!" And Feynman pointed out: "It [renormalization] is still what I would call a dippy process! Having to resort to such hocus-pocus has prevented us from proving that the theory of quantum electrodynamics is mathematically self-consistent. … I suspect that renormalization is not mathematically legitimate.”

One can debate if the avoidance of divergences by renormalization is a more or less fruitful concept or a fundamental principle.

In any case, meanwhile there are new perspectives to renormalization. For instance, one of these approaches, instead of being concerned with good limiting behavior as the cut-off for great momenta, focuses on the behavior of the running constants, i.e., how e.g. charge or mass vary with increasing energy (i.e., decreasing distance).

Of course, in our short introduction we have to leave many questions open, for instance: Are there other theories besides QED which are renormalizable? Are there theories which are not renormalizable? What is the crucial property for a theory to be renormalizable?, to mention a few. Another interesting point would be the discussion of Haag’s theorem which states that the interaction picture of a relativistic quantum field theory does not exist, thus shaking the foundations of QFT.

Basically, the issue expresses our ignorance of physics at very small distances. Is the electron really a mass point? Experimentally, this property could only be confirmed up to distances of about 10^{-18} m. But it is still an open question how the micro-world behaves at distances of say the Planck length 10^{-33} m. At this scale, there may be effects, still unknown to us, which, in principle, could repair the divergences.

Appendix X
Exercises and Solutions

X.1 Exercises, Chap. 15

1. Given the potential step

$$V(x) = \begin{cases} 0 & \text{for} \quad x > 0 \\ V_0 > 0 & \quad x \le 0. \end{cases} \tag{X.1}$$

The incident quantum object is described as a plane wave running from the right to the left with $E > V_0$. Determine the transmission and reflection coefficients.

2. Given a finite potential well of depth V_0 and width L; estimate the number of energy levels.
Solution: We have the inequality (15.43):

$$V_0 - \frac{\hbar^2}{2m}\left(\frac{N}{2L}\pi\right)^2 < |E| < V_0 - \frac{\hbar^2}{2m}\left(\frac{N-1}{2L}\pi\right)^2. \tag{X.2}$$

If N_0, but not N_0+ levels exist, this means that

$$V_0 - \frac{\hbar^2}{2m}\left(\frac{N_0-1}{2L}\pi\right)^2 > 0 \text{ and } V_0 - \frac{\hbar^2}{2m}\left(\frac{N_0}{2L}\pi\right)^2 < 0, \tag{X.3}$$

i.e.

$$N_0 - 1 < \frac{2L}{\pi\hbar}\sqrt{2mV_0} < N_0. \tag{X.4}$$

Solving for N_0, we find

$$\frac{2L}{\pi\hbar}\sqrt{2mV_0} < N_0 < 1 + \frac{2L}{\pi\hbar}\sqrt{2mV_0}. \tag{X.5}$$

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

3. Given a delta potential at $x = 0$; determine the spectrum (negative potential, $E < 0$) and the situation for scattering (positive potential, $E > 0$).
Solution: The SEq reads

$$E\varphi = -\frac{\hbar^2}{2m}\varphi'' + V\delta(x)\varphi; \; V = \frac{\hbar^2}{2m}\gamma. \tag{X.6}$$

We integrate this equation between $-\varepsilon$ and ε, where ε is a small parameter which we subsequently allow to go to zero. First, it follows that

$$E\int_{-\varepsilon}^{\varepsilon}\varphi \mathrm{d}x = -\frac{\hbar^2}{2m}\left[\varphi'(\varepsilon) - \varphi'(-\varepsilon)\right] + V\varphi(0). \tag{X.7}$$

We assume that φ is continuous at the origin, but the derivative can have a discontinuity (as for the infinite potential wall). It follows for $\varepsilon \to 0$ that:

$$0 = -\frac{\hbar^2}{2m}\left[\varphi'(+0) - \varphi'(-0)\right] + V\varphi(0) \tag{X.8}$$

or

$$\varphi'(+0) - \varphi'(-0) = \frac{2m}{\hbar^2}V\varphi(0) = \gamma\varphi(0). \tag{X.9}$$

Concrete forms (bound or free) for the wavefunction have to be inserted into this equation.

(a) Bound case, i.e. $V < 0$: With

$$\varphi(x < 0) = Ae^{\kappa x}; \; \varphi(x >) = Be^{-\kappa x}, \tag{X.10}$$

it follows that

$$A = B; \; -\kappa B - \kappa A = \gamma A \tag{X.11}$$

and thus

$$\kappa = -\frac{1}{2}\gamma A. \tag{X.12}$$

That is, for $V < 0$ there is always exactly one bound state.

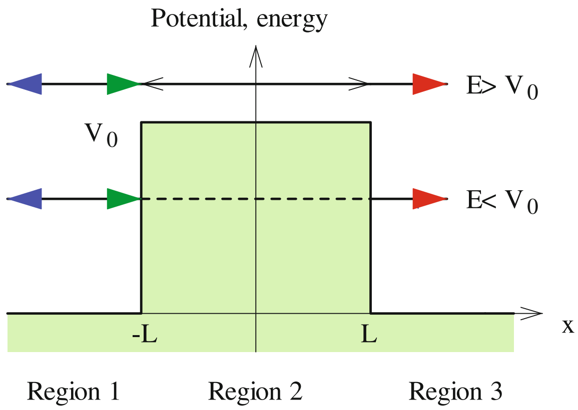
(b) Scattering, i.e. $V > 0$: The quantum object comes from the left. Then we have

$$\varphi(x < 0) = Ae^{ikx} + Be^{-ikx}; \; \varphi(x > 0) = Ce^{ikx}. \tag{X.13}$$

This leads to

$$A + B = C; \; ikC - (ikA - ikB) = \gamma C \tag{X.14}$$

Fig. X.1 Scattering at the potential barrier



and it follows that

$$B = \frac{\gamma}{2ik - \gamma}A; \; C = \frac{2ik}{2ik - \gamma}A. \tag{X.15}$$

This yields for the transmission and reflection coefficients:

$$T = \frac{4k^2}{4k^2 + \gamma^2}; \; R = \frac{\gamma^2}{4k^2 + \gamma^2}. \tag{X.16}$$

4. Given the potential barrier

$$V(x) = \begin{cases} V_0 > 0 & \text{for} \quad -L < x < L \\ 0 & \quad\quad\;\; \text{otherwise.} \end{cases} \tag{X.17}$$

The incident quantum object is described by a plane wave running from the left to the right. Determine the transmission and reflection coefficients.
Solution: The potential is outlined in Fig. X.1.
We treat the cases $E > V_0$ and $E < V_0$ together by setting

$$\gamma = \begin{cases} \kappa \\ ik' \end{cases} \text{for} \quad \begin{matrix} E < V_0 \\ E > V_0 \end{matrix} \quad \text{with} \quad \begin{matrix} \kappa^2 = \frac{2m}{\hbar^2}(V_0 - E) \\ k'^2 = \frac{2m}{\hbar^2}(E - V_0). \end{matrix} \tag{X.18}$$

The solutions in the different regions are

$$\begin{aligned} \varphi_1 &= Ae^{ikx} + Be^{-ikx} \\ \varphi_2 &= Ce^{\gamma x} + De^{-\gamma x} \\ \varphi_3 &= Fe^{ikx} + Ge^{-ikx} \end{aligned} \tag{X.19}$$

with $k^2 = \frac{2m}{\hbar^2}E$.

Determination of the Integration Constant:

The incoming wave comes from the left with the amplitude A; thus, in region 3 there is no wave running from the right to the left and therefore we have $G = 0$. At the discontinuities $x = \pm L$, we have

$$\begin{aligned} Ae^{-ikL} + Be^{ikL} &= Ce^{-\gamma L} + De^{\gamma L} \\ ikAe^{-ikL} - ikBe^{ikL} &= \gamma Ce^{-\gamma L} - \gamma De^{\gamma L} \end{aligned} \tag{X.20}$$

and

$$\begin{aligned} Ce^{\gamma L} + De^{-\gamma L} &= Fe^{ikL} \\ \gamma Ce^{\gamma L} - \gamma De^{-\gamma L} &= ikFe^{ikL} \end{aligned} \tag{X.21}$$

From (X.20), it follows that

$$2ikAe^{-ikL} = (\gamma + ik)\, Ce^{-\gamma L} - (\gamma - ik)\, De^{\gamma L}, \tag{X.22}$$

and from (X.21) we have:

$$(\gamma - ik)\, Ce^{\gamma L} - (\gamma + ik)\, De^{-\gamma L} = 0. \tag{X.23}$$

This leads directly to

$$C = \frac{\gamma + ik}{\gamma - ik} De^{-2\gamma L}. \tag{X.24}$$

Insertion yields

$$\begin{aligned} 2ikAe^{-ikL} &= (\gamma + ik)\, \frac{(\gamma + ik)}{(\gamma - ik)} De^{-2\gamma L} e^{-\gamma L} - (\gamma - ik)\, De^{\gamma L} \\ \rightarrow 2ikAe^{-ikL} &= \frac{e^{-2\gamma L} (\gamma + ik)^2 - e^{2\gamma L} (\gamma - ik)^2}{(\gamma - ik)} De^{-\gamma L}. \end{aligned} \tag{X.25}$$

We introduce the shorthand notation

$$M = e^{-2\gamma L} (\gamma + ik)^2 - e^{2\gamma L} (\gamma - ik)^2 . \tag{X.26}$$

We then have

$$\begin{aligned} 2ikAe^{-ikL} &= \frac{M}{(\gamma - ik)} De^{-\gamma L} \rightarrow \\ D &= \frac{(\gamma - ik)}{M} 2ikAe^{-ikL} e^{\gamma L} = \frac{2ik\, (\gamma - ik)}{M} Ae^{-ikL} e^{\gamma L} \end{aligned} \tag{X.27}$$

and from this with (X.24)

$$C = \frac{(\gamma + ik)}{(\gamma - ik)} De^{-2\gamma L} = \frac{(\gamma + ik)}{(\gamma - ik)} \frac{2ik(\gamma - ik)}{M} Ae^{-ikL} e^{\gamma L} e^{-2\gamma L} \\ = \frac{2ik(\gamma + ik)}{M} Ae^{-ikL} e^{-\gamma L}. \tag{X.28}$$

Due to

$$Be^{ikL} = Ce^{-\gamma L} + De^{\gamma L} - Ae^{-ikL}; \tag{X.29}$$

this leads for the constant B to

$$\begin{aligned} B &= \frac{2ik(\gamma + ik)e^{-2\gamma L} + 2ik(\gamma - ik)e^{2\gamma L} - M}{M} Ae^{-2ikL} \\ &= \frac{[2ik - \gamma - ik](\gamma + ik)e^{-2\gamma L} + [2ik + \gamma - ik](\gamma - ik)e^{2\gamma L}}{M} Ae^{-2ikL} \\ &= \frac{e^{2\gamma L} - e^{-2\gamma L}}{M} \left(\gamma^2 + k^2\right) Ae^{-2ikL}, \end{aligned} \tag{X.30}$$

and due to

$$Fe^{ikL} = Ce^{\gamma L} + De^{-\gamma L}, \tag{X.31}$$

we find for the constant F:

$$F = \frac{2ik(\gamma + ik)}{M} Ae^{-2ikL} e^{-\gamma L} e^{\gamma L} + \frac{2ik(\gamma - ik)}{M} Ae^{-2ikL} e^{\gamma L} e^{-\gamma L} \\ = \frac{4ik\gamma}{M} Ae^{-2ikL}. \tag{X.32}$$

In sum

$$B = \frac{e^{2\gamma L} - e^{-2\gamma L}}{M} \left(\gamma^2 + k^2\right) Ae^{-2ikL} \\ F = \frac{4ik\gamma}{M} Ae^{-2ikL}, \tag{X.33}$$

where it holds that

$$\gamma = \begin{cases} \kappa & \text{for} \quad E < V_0 \\ ik' & \phantom{\text{for}} \quad E > V_0 \end{cases} \tag{X.34}$$

$$M = \begin{cases} e^{-2\kappa L}(\kappa + ik)^2 - e^{2\kappa L}(\kappa - ik)^2 \\ e^{-2ik'L}\left(ik' + ik\right)^2 - e^{2ik'L}\left(ik' - ik\right)^2. \end{cases} \tag{X.35}$$

Determination of T and R:

The partial waves of interest are

$$\varphi_{\text{ein}} = Ae^{ikx};\ \varphi_{\text{refl}} = Be^{-ikx};\ \varphi_{\text{trans}} = Fe^{ikx}. \tag{X.36}$$

We have

$$T = \frac{|F|^2}{|A|^2} = \frac{\left|\frac{4ik\gamma}{M}Ae^{-2ikL}\right|^2}{|A|^2} = \left|\frac{4k\gamma}{M}\right|^2 = \frac{16k^2\gamma\gamma^*}{MM^*} \tag{X.37}$$

and

$$\begin{aligned} R = \frac{|B|^2}{|A|^2} &= \frac{\left|\frac{e^{2\gamma L}-e^{-2\gamma L}}{M}\left(\gamma^2 + k^2\right)Ae^{-2ikL}\right|^2}{|A|^2} \\ &= \frac{\left(e^{2\gamma L} - e^{-2\gamma L}\right)\left(e^{2\gamma^* L} - e^{-2\gamma^* L}\right)\left(\gamma^2 + k^2\right)\left(\gamma^{*2} + k^2\right)}{MM^*}. \end{aligned} \tag{X.38}$$

We confine the discussion to T. We have:

$$T = \frac{16k^2\gamma\gamma^*}{MM^*} = \begin{cases} \dfrac{16k^2\kappa^2}{MM^*} & \text{for } E < V_0 \\ \dfrac{16k^2k'^2}{MM^*} & \text{for } E > V_0 \end{cases} \tag{X.39}$$

with

$$MM^* = \begin{cases} \left[e^{-2\kappa L}(\kappa + ik)^2 - e^{2\kappa L}(\kappa - ik)^2\right] \\ \left[e^{-2\kappa L}(\kappa - ik)^2 - e^{2\kappa L}(\kappa + ik)^2\right] \\ \left[-e^{-2ik'L}\left(k' + k\right)^2 + e^{2ik'L}\left(k' - k\right)^2\right] \\ \left[-e^{2ik'L}\left(k' + k\right)^2 + e^{-2ik'L}\left(k' - k\right)^2\right] \end{cases} \tag{X.40}$$

or

$$MM^* = \begin{cases} 2\cosh 4\kappa L \cdot (\kappa - ik)^2(\kappa + ik)^2 - (\kappa + ik)^2(\kappa + ik)^2 \\ \quad - (\kappa - ik)^2(\kappa - ik)^2 \\ \left(k' + k\right)^4 + \left(k' - k\right)^4 - 2\cos 4ik'L \cdot \left(k' - k\right)^2\left(k' + k\right)^2 \end{cases} \tag{X.41}$$

or

$$MM^* = \begin{cases} 2\left(\kappa^2 + k^2\right)^2\cosh 4\kappa L - 2\left[\kappa^4 - 6\kappa^2k^2 + k^4\right] \\ 2\left[k'^4 + 6k'^2k^2 + k^4\right] - 2\left(k'^2 - k^2\right)^2\cos 4k'L. \end{cases} \tag{X.42}$$

This gives

$$T = \begin{cases} \frac{16k^2\kappa^2}{MM^*} = \frac{8k^2\kappa^2}{\left(\kappa^2 + k^2\right)^2 \cosh 4\kappa L - \left(\kappa^4 - 6\kappa^2 k^2 + k^4\right)} \\ \frac{16k^2 k'^2}{MM^*} = \frac{8k^2 k'^2}{\left(k'^4 + 6k'^2 k^2 + k^4\right) - \left(k'^2 - k^2\right)^2 \cos 4k'L} \end{cases} \text{for } \begin{matrix} E < V_0 \\ E > V_0. \end{matrix} \tag{X.43}$$

Finally, we insert E and V_0; with

$$k^2 = \frac{2m}{\hbar^2} E;\ k'^2 = \frac{2m}{\hbar^2} (E - V_0);\ \kappa^2 = \frac{2m}{\hbar^2} (V_0 - E) \tag{X.44}$$

we obtain

$$T = \begin{cases} \frac{8E (E - V_0)}{8E (E - V_0) + V_0^2 (1 - \cosh 4\kappa L)} \\ \frac{8E (E - V_0)}{8E (E - V_0) + V_0^2 (1 - \cos 4k'L)} \end{cases} \text{for } \begin{matrix} E < V_0 \\ E > V_0. \end{matrix} \tag{X.45}$$

In order to write this more compactly, we introduce the abbreviations

$$z = \frac{E}{V_0};\ \mu = \sqrt{\frac{2m}{\hbar^2} V_0 L^2}. \tag{X.46}$$

With

$$k'L = \mu\sqrt{z - 1};\ \kappa L = \mu\sqrt{1 - z} \tag{X.47}$$

it follows finally that

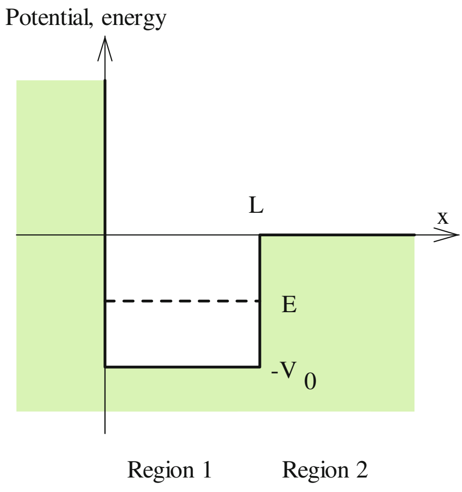
$$T = \begin{cases} \frac{8z (z - 1)}{8z (z - 1) + 1 - \cosh 4\mu\sqrt{1 - z}} \\ \frac{8z (z - 1)}{8z (z - 1) + 1 - \cos 4\mu\sqrt{z - 1}} \end{cases} \text{for } \begin{matrix} E < V_0;\ 0 < z < 1 \\ E > V_0;\ z > 1 \end{matrix} \tag{X.48}$$

or, with $\cos ix = \cosh x$ and in one formula:

$$T = \frac{8z (z - 1)}{8z (z - 1) + 1 - \cosh 4\mu\sqrt{1 - z}} \text{ for } 0 < z = \frac{E}{V_0}. \tag{X.49}$$

The graphical representation of T as a function of $z = \frac{E}{V_0}$ is found in Chap. 15.

Fig. X.2 One-sided infinite potential well



5. Given the one-sided infinite potential well

$$V(x) = \begin{cases} 0 & & L < x \\ -V_0 & \text{for} & 0 < x \leq L \\ \infty & & x \leq 0 \end{cases} \tag{X.50}$$

with $V_0 > 0$. For the energy, let $-V_0 < E < 0$. Sketch the potential. Determine the stationary SEq in the different regions and deduce from them an *ansatz* for the wavefunction. Adjust the wavefunctions at the discontinuities and show that the allowed energy levels are defined by the equation $k \cot kL = -\kappa$ with $k^2 = 2m(V_0 + E)/\hbar^2$ and $\kappa^2 = -2mE/\hbar^2$. Is there always (i.e. for all V_0) a bound state?
Solution: The potential is outlined in Fig. X.2.
The stationary SEq are

$$E\varphi_1 = -\frac{\hbar^2}{2m}\varphi_1'' - V_0\varphi_1;\ E\varphi_2 = -\frac{\hbar^2}{2m}\varphi_2''. \tag{X.51}$$

They have the solutions

$$\begin{array}{c} \varphi_1 = Ae^{ikx} + Be^{-ikx};\ \varphi_2 = Ce^{\kappa x} + De^{-\kappa x} \\ k^2 = 2m(V_0 - |E|)/\hbar^2;\ \kappa^2 = 2m|E|/\hbar^2. \end{array} \tag{X.52}$$

For $x < 0$, $\varphi \equiv 0$.

Hence, the matching conditions at $x = 0$ are[157] $\varphi_1(0) = 0$. For $x > 0$, C must vanish since otherwise the solution is not bounded. So we find

$$\begin{array}{ll} x = 0: & B = -A \\ x = L: & \begin{array}{c} Ae^{ikL} - ABe^{-ikL} = 2iA\sin kL = De^{-\kappa L} \\ ikAe^{ikL} + ikAe^{-ikL} = 2ikA\cos kL = -\kappa De^{-\kappa L}. \end{array} \end{array} \tag{X.53}$$

Division of the last two equations gives

$$k\cot kL = -\kappa \text{ or } \tan kL = -\frac{k}{\kappa}. \tag{X.54}$$

This is the quantization condition for the energy. The equation is not solvable in closed form, but with the following considerations we can obtain some more information.
The tangent is periodic in π; thus it holds that

$$\tan(kL + m\pi) = -\frac{k}{\kappa};\ m = 0, \pm1, \pm2, \ldots \tag{X.55}$$

Since k, κ and L are positive, we can rewrite this equation as

$$kL = n\pi - \arctan\frac{k}{\kappa};\ n = 1, 2, \ldots \tag{X.56}$$

We rewrite this again, as in Chap. 15, as an inequality, making use of $0 < \arctan x < \frac{\pi}{2}$ (due to $x > 0$). Then it follows initially that:

$$\frac{2n-1}{2}\pi < kL < n\pi. \tag{X.57}$$

Each of these three terms is positive; so we can square and insert the relation $k^2 = 2m(V_0 - |E|)/\hbar^2$. Resorting and solving for $|E|$, we find:

$$V_0 - \frac{1}{2m}\left(n\frac{\hbar\pi}{L}\right)^2 < |E| < V_0 - \frac{1}{2m}\left(\frac{2n-1}{2}\frac{\hbar\pi}{L}\right)^2. \tag{X.58}$$

In a finite potential well, there is always at least one bound state (see Chap. 15); does this apply here, too? We set $n = 1$ and obtain

$$V_0 - \frac{1}{2m}\left(\frac{\hbar\pi}{L}\right)^2 < |E| < V_0 - \frac{1}{2m}\left(\frac{\hbar\pi}{2L}\right)^2. \tag{X.59}$$

Hence, a bound state can exist only for

[157] Note: Due to the infinite jump of the potential, only the wavefunction can be matched at $x = 0$, but not the derivative.

$$V_0 > \frac{1}{2m}\left(\frac{\hbar\pi}{2L}\right)^2. \tag{X.60}$$

Finally, we ask for the conditions that there are exactly N energy levels. This is the case if the right side of (X.58) is positive for $n = N$, but negative for $n = N + 1$, i.e. if

$$V_0 - \frac{1}{2m}\left(\frac{2N-1}{2}\frac{\hbar\pi}{L}\right)^2 > 0;\; V_0 - \frac{1}{2m}\left(\frac{2N+1}{2}\frac{\hbar\pi}{L}\right)^2 < 0. \tag{X.61}$$

It follows that

$$\left(N - \frac{1}{2}\right)^2 < \frac{2mL^2}{\hbar^2\pi^2}V_0 < \left(N + \frac{1}{2}\right)^2. \tag{X.62}$$

6. Given the potential

$$V(x) = \begin{cases} \infty & & x < 0 \\ V_0 > 0 & \text{for} & 0 \le x \le L \\ 0 & & L < x \end{cases}. \tag{X.63}$$

An object described by a plane wave passes from the right towards the origin. Sketch the potential. Calculate the wavefunction for the case $E < V_0$. Which regions are classically allowed, which are not? Determine first the stationary SEq's in the different regions and solve them with an appropriate *ansatz*. Are all the mathematical solutions physically allowed? Determine the free constants using the continuity conditions at the discontinuities of the potential.
Perform the calculations for the case $E > V_0$, also.

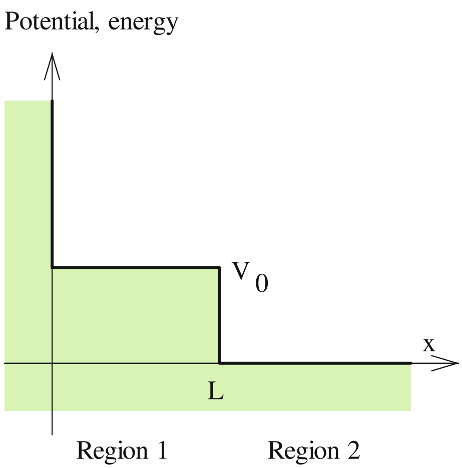


Fig. X.3 The potential (X.63)

Solution: The potential is outlined in Fig. X.3. The region $x < 0$ as well as region 1 for $E < V_0$ are classically forbidden; region 2 as well as region 1 for $E > V_0$ are classically allowed. The SEq in regions 1 and 2 are

$$E\varphi_1 = -\frac{\hbar^2}{2m}\varphi_1'' + V_0\varphi_1;\; E\varphi_2 = -\frac{\hbar^2}{2m}\varphi_2'', \tag{X.64}$$

and it follows that

$$\varphi_1'' = -\frac{2m}{\hbar^2}(E - V_0)\,\varphi_1 = \gamma^2\varphi_1;\; \varphi_2'' = -\frac{2m}{\hbar^2}E\varphi_2 = -k^2\varphi_2 \tag{X.65}$$

with

$$\begin{matrix} \gamma^2 = -k'^2 \\ \gamma^2 = \kappa^2 \end{matrix} \quad \text{for} \quad \begin{matrix} E > V_0 \\ E < V_0 \end{matrix} \quad \text{or} \quad \begin{matrix} k'^2 = \frac{2m}{\hbar^2}(E - V_0) \\ \kappa^2 = \frac{2m}{\hbar^2}(V_0 - E)\,. \end{matrix} \tag{X.66}$$

The solutions are

$$\varphi_1 = Ce^{\gamma x} + De^{-\gamma x};\; \varphi_2 = Fe^{ikx} + Ge^{-ikx}. \tag{X.67}$$

All four partial solutions are physically allowed (i.e. all four integration constants are nonzero). In particular, we have

$$\varphi_{\text{ein}} = Ge^{-ikx};\; \varphi_{\text{refl}} = Fe^{ikx}. \tag{X.68}$$

The matching conditions

$$\varphi_1(0) = 0;\; \varphi_1(L) = \varphi_2(L)\,;\; \varphi_1'(L) = \varphi_2'(L) \tag{X.69}$$

yield

$$\begin{aligned} D = -C;\; Ce^{\gamma L} + De^{-\gamma L} &= Fe^{ikL} + Ge^{-ikL} \\ \gamma Ce^{\gamma L} - \gamma De^{-\gamma L} &= ikFe^{ikL} - ikGe^{-ikL}. \end{aligned} \tag{X.70}$$

Resolving gives

$$\begin{aligned} C &= \frac{2k}{e^{\gamma L}(k + i\gamma) - e^{-\gamma L}(k - i\gamma)}Ge^{-ikL} \\ F &= \frac{e^{\gamma L}(k - i\gamma) - e^{-\gamma L}(k + i\gamma)}{e^{\gamma L}(k + i\gamma) - e^{-\gamma L}(k - i\gamma)}Ge^{-2ikL}. \end{aligned} \tag{X.71}$$

We calculate the reflection coefficient R by using the probability current density:

$$j_{\text{ein}} = \frac{\hbar k}{m}|G|^2\,;\; j_{\text{refl}} = \frac{\hbar k}{m}|F|^2\;;\; R = \left|\frac{j_{\text{refl}}}{j_{\text{ein}}}\right| = \left|\frac{F}{G}\right|^2. \tag{X.72}$$

Since R depends only on the term F/G, we consider in the following only F. For simplicity we rearrange:

$$F = \frac{k \sinh \gamma L - i\gamma \cosh \gamma L}{k \sinh \gamma L + i\gamma \cosh \gamma L} G e^{-2ikL} = -\frac{\cosh \gamma L + i\frac{k}{\gamma} \sinh \gamma}{\cosh \gamma L - i\frac{k}{\gamma} \sinh \gamma} G e^{-2ikL}. \tag{X.73}$$

Because of $\cosh(ix) = \cos x$ and $\sinh(ix) = \sin x$, both terms $\cosh \gamma L$ and $\frac{k}{\gamma} \sinh \gamma$ are real for $\gamma = \kappa$ and for $\gamma = ik'$. This means that

$$\cosh \gamma L + i\frac{k}{\gamma} \sinh \gamma = \sqrt{\cosh^2 \gamma L + \left(\frac{k}{\gamma}\right)^2 \sinh^2 \gamma} \cdot e^{i \arctan \frac{k}{\gamma} \tanh \gamma}, \tag{X.74}$$

and from this, it follows that

$$F = -e^{2i \arctan \frac{k}{\gamma} \tanh \gamma} e^{-2ikL} G. \tag{X.75}$$

Since $\frac{k}{\gamma} \tanh \gamma$ is real for $\gamma = \kappa$ and for $\gamma = ik'$, we see directly and without much arithmetic that for $E < V_0$ as well as for $E > V_0$, we have:

$$R = \left|\frac{F}{G}\right|^2 = 1 \tag{X.76}$$

as indeed must hold.

7. Given a potential step embedded in an infinite potential well

$$V(x) = \begin{cases} 0 & & 0 < x < L \\ V_0 > 0 & \text{for} & -L < x \leq 0 \\ \infty & & x \geq |L| \end{cases}. \tag{X.77}$$

Calculate the spectrum for $E > V_0$.
Solution: The potential is outlined in Fig. X.4. The wavefunctions are given by

$$\varphi_1 = Ae^{ik_1x} + Be^{-ik_1x}; \ \varphi_2 = Ce^{ik_2x} + De^{-ik_2x} \tag{X.78}$$

with

$$k_1^2 = \frac{2m}{\hbar^2}(E - V_0) = k_2^2 - \frac{2m}{\hbar^2}V_0; \ k_2^2 = \frac{2m}{\hbar^2}E. \tag{X.79}$$

At $x = 0$, we find:

$$\begin{aligned} A + B &= C + D \\ ik_1A - ik_1B &= ik_2C - ik_2D \end{aligned} \tag{X.80}$$

and at $x = \pm L$:

$$\begin{aligned} Ae^{-ik_1L} + Be^{ik_1L} &= 0 \\ Ce^{ik_2L} + De^{-ik_2L} &= 0. \end{aligned} \tag{X.81}$$

These are four equations for four unknowns; in order that a solution exists, the determinant of the coefficients D must vanish. This means that

$$D = \begin{vmatrix} 1 & 1 & -1 & -1 \\ k_1 & -k_1 & -k_2 & k_2 \\ e^{-ik_1L} & e^{ik_1L} & 0 & 0 \\ 0 & 0 & e^{ik_2L} & e^{-ik_2L} \end{vmatrix} = 0. \tag{X.82}$$

The determinant is calculated to give (e.g. by expansion with respect to the third row):

$$D = 2i\left[(k_2 + k_1)\sin(k_2L + k_1L) - (k_2 - k_1)\sin(k_2L - k_1L)\right]. \tag{X.83}$$

Hence, the energy levels are determined by the equation

$$\sin(k_2L + k_1L) = \frac{k_2 - k_1}{k_2 + k_1}\sin(k_2L - k_1L). \tag{X.84}$$

A closed solution does not exist; however, we can see directly from (X.84) that the spectrum is discrete (of course, this is due to the infinite potential walls). In the following, we are concerned only with the approximate solution of (X.84). First, we note that the ratio on the right side goes to zero for large E, since with (X.79) we have:

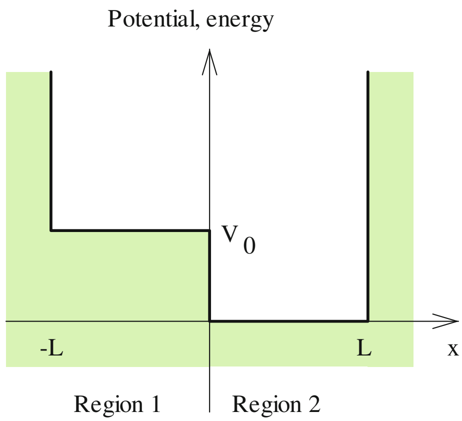


Fig. X.4 The potential of (X.77)

$$\frac{k_2 - k_1}{k_2 + k_1} = \frac{\sqrt{E} - \sqrt{E - V_0}}{\sqrt{E} + \sqrt{E - V_0}} = \frac{1 - \sqrt{1 - \frac{V_0}{E}}}{1 + \sqrt{1 - \frac{V_0}{E}}} \approx \frac{V_0}{4E}. \tag{X.85}$$

Hence, for sufficiently large E, the right side of (X.84) is approximately zero. Then it follows that

$$(k_2 + k_1) L = N\pi + \varepsilon_N, \tag{X.86}$$

where ε_N is a small correction term and N a sufficiently large natural number (because of $k_2 + k_1 \approx 2\sqrt{\frac{2m}{\hbar^2}E}$).
We rewrite (X.84) in the form

$$\sin\left((k_2 + k_1) L\right) = \frac{L^2\left(k_2^2 - k_1^2\right)}{L^2\left(k_2 + k_1\right)^2} \sin\left(\frac{L^2\left(k_2^2 - k_1^2\right)}{L\left(k_2 + k_1\right)}\right). \tag{X.87}$$

With (X.79), it follows that:

$$L^2\left(k_2^2 - k_1^2\right) = L^2 \frac{2m}{\hbar^2} V_0 = \mu^2 \tag{X.88}$$

and this gives

$$\sin\left((k_2 + k_1) L\right) = \frac{\mu^2}{L^2\left(k_2 + k_1\right)^2} \sin\left(\frac{\mu^2}{L\left(k_2 + k_1\right)}\right); \tag{X.89}$$

and with (X.86)

$$\sin\left(N\pi + \varepsilon_N\right) = \frac{\mu^2}{\left(N\pi + \varepsilon_N\right)^2} \sin\left(\frac{\mu^2}{N\pi + \varepsilon_N}\right). \tag{X.90}$$

The left side equals $(-1)^N \sin\varepsilon_N$. For sufficiently large energies (which corresponds to large N and small ε_N), we can use the approximation $\sin x \approx x$ and obtain the approximate result:

$$(-1)^N \varepsilon_N = \frac{\mu^4}{\left(N\pi + \varepsilon_N\right)^3} \text{ or } \varepsilon_N = (-1)^N \frac{\mu^4}{(N\pi)^3} \sim \frac{1}{N^3}. \tag{X.91}$$

The energy levels follow from (X.86). We notice that due to (X.79), it holds generally that

$$k_2 + k_1 = \alpha \rightarrow E = \frac{V_0}{4}\left(\sqrt{\frac{\hbar^2}{2mV_0}}\alpha + \sqrt{\frac{2mV_0}{\hbar^2}}\frac{1}{\alpha}\right)^2$$
$$= \frac{V_0}{4}\left(\frac{\alpha}{\mu L} + \frac{\mu L}{\alpha}\right)^2. \qquad \text{(X.92)}$$

With $\alpha = \frac{N\pi + \varepsilon_N}{L}$, it follows that

$$E_N = \frac{V_0}{4}\left(\frac{N\pi + \varepsilon_N}{\mu} + \frac{\mu}{N\pi + \varepsilon_N}\right)^2. \qquad \text{(X.93)}$$

We expand the right side in terms of powers of N and keep only the two largest terms. Due to $\varepsilon_N \sim N^{-3}$, we can neglect the correction term ε_N, and obtain approximately

$$E_N \approx \frac{V_0}{4}\frac{N^2\pi^2}{\mu^2}\left(1 + 2\frac{\mu^2}{N^2\pi^2}\right) = \frac{\hbar^2 N^2\pi^2}{8mL^2}\left(1 + \frac{4mV_0L^2}{N^2\pi^2\hbar^2}\right). \qquad \text{(X.94)}$$

So we have a discrete energy spectrum which for sufficiently large N (i.e. for sufficiently high energies) is by and large that of the infinite potential well. The existence of the potential step results essentially in a slight raising of all the energy levels.

8. (Resonances) Given a potential barrier in front of an infinite potential wall:

$$V(x) = \begin{cases} \infty & x < 0 \\ V_0 > 0 \quad \text{for} & a \le x \le b. \\ 0 & \text{otherwise}. \end{cases} \qquad \text{(X.95)}$$

The incident quantum object has the energy $E < V_0$ and comes from the right. For which parameter values is the phase shift of the outgoing wave particularly large/does the phase change especially fast? What is the physical explanation?
Solution: The potential is outlined in Fig. X.5.
The *ansatz* is (region 1: $0 \le x \le a$; region 2: $a \le x \le b$; region 3: $b \le x$)

$$\begin{aligned} \psi_1 &= Ae^{ikx} + Be^{-ikx} \\ \psi_2 &= Ce^{\kappa x} + De^{-\kappa x} \\ \psi_3 &= Fe^{ikx} + Ge^{-ikx} \end{aligned} \qquad \text{(X.96)}$$

with $k^2 = 2mE/\hbar^2$ and $\kappa^2 = 2m(V_0 - E)/\hbar^2$. The term Ge^{-ikx} describes the incoming and Fe^{ikx} the scattered (reflected) object.
At $x = 0$, we have $A = -B$. For the two other discontinuities, we find

$$\begin{aligned} Ae^{ika} - Ae^{-ika} &= Ce^{\kappa a} + De^{-\kappa a} \\ ikAe^{ika} + ikAe^{-ika} &= \kappa Ce^{\kappa a} - \kappa De^{-\kappa a} \end{aligned} \tag{X.97}$$

and

$$\begin{aligned} Ce^{\kappa b} + De^{-\kappa b} &= Fe^{ikb} + Ge^{-ikb} \\ \kappa Ce^{\kappa b} - \kappa De^{-\kappa b} &= ikFe^{ikb} - ikGe^{-ikb}. \end{aligned} \tag{X.98}$$

With the abbreviations

$$\begin{aligned} A' &= 2iA;\; C' = Ce^{\kappa a};\; D' = De^{-\kappa a} \\ F' &= Fe^{ikb};\; G' = Ge^{-ikb};\; d = b - a, \end{aligned} \tag{X.99}$$

it follows that

$$\begin{aligned} A' \sin ka &= C' + D' \\ \frac{k}{\kappa} A' \cos ka &= C' - D' \end{aligned} \tag{X.100}$$

and

$$\begin{aligned} C'e^{\kappa d} + D'e^{-\kappa d} &= F' + G' \\ \kappa C'e^{\kappa d} - \kappa D'e^{-\kappa d} &= ikF' - ikG'. \end{aligned} \tag{X.101}$$

This leads initially to

$$C' = A' \frac{\sin ka + \frac{k}{\kappa} \cos ka}{2};\; D' = A' \frac{\sin ka - \frac{k}{\kappa} \cos ka}{2} \tag{X.102}$$

and thus to

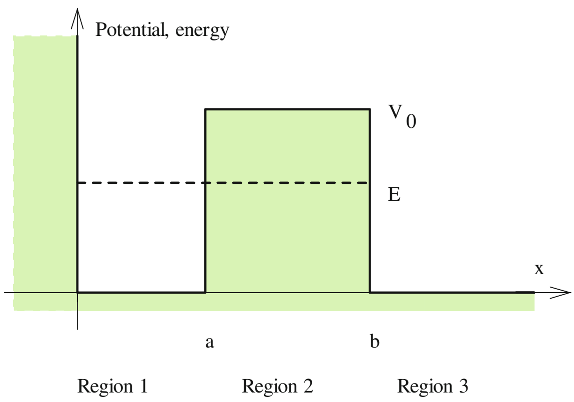


Fig. X.5 The potential of (X.95)

$$A'\frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}e^{\kappa d} + A'\frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}e^{-\kappa d} = F' + G'$$
$$\kappa A'\frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}e^{\kappa d} - \kappa A'\frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}e^{-\kappa d} = ikF' - ikG'. \tag{X.103}$$

It follows that

$$A'\frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}e^{\kappa d} + A'\frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}e^{-\kappa d} = F' + G'$$
$$A'\frac{\kappa}{ik}\frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}e^{\kappa d} - A'\frac{\kappa}{ik}\frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}e^{-\kappa d} = F' - G' \tag{X.104}$$

and therefore

$$\begin{aligned}2\frac{F'}{A'} &= \frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}e^{\kappa d} + \frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}e^{-\kappa d}\\ &\quad + \frac{\kappa}{ik}\frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}e^{\kappa d} - \frac{\kappa}{ik}\frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}e^{-\kappa d}\\ 2\frac{G'}{A'} &= \frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}e^{\kappa d} + \frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}e^{-\kappa d}\\ &\quad - \frac{\kappa}{ik}\frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}e^{\kappa d} + \frac{\kappa}{ik}\frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}e^{-\kappa d}\end{aligned} \tag{X.105}$$

or

$$2\frac{F'}{A'} = \frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}\left(1 + \frac{\kappa}{ik}\right)e^{\kappa d} + \frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}\left(1 - \frac{\kappa}{ik}\right)e^{-\kappa d}$$
$$2\frac{G'}{A'} = \frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}\left(1 - \frac{\kappa}{ik}\right)e^{\kappa d} + \frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}\left(1 + \frac{\kappa}{ik}\right)e^{-\kappa d} \tag{X.106}$$

and finally

$$\frac{F'}{G'} = \frac{\frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}\left(1 + \frac{\kappa}{ik}\right)e^{\kappa d} + \frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}\left(1 - \frac{\kappa}{ik}\right)e^{-\kappa d}}{\frac{\sin ka + \frac{k}{\kappa}\cos ka}{2}\left(1 - \frac{\kappa}{ik}\right)e^{\kappa d} + \frac{\sin ka - \frac{k}{\kappa}\cos ka}{2}\left(1 + \frac{\kappa}{ik}\right)e^{-\kappa d}}. \tag{X.107}$$

We see directly that the right side has the form $\frac{z^*}{z}$ and therefore the absolute value 1 (as it must be; what comes in goes out again).
We rewrite the result:

$$\frac{F'}{G'} = \frac{(\kappa + k\cot ka)\left(1 + \frac{\kappa}{ik}\right)e^{\kappa d} + (\kappa - k\cot ka)\left(1 - \frac{\kappa}{ik}\right)e^{-\kappa d}}{(\kappa + k\cot ka)\left(1 - \frac{\kappa}{ik}\right)e^{\kappa d} + (\kappa - k\cot ka)\left(1 + \frac{\kappa}{ik}\right)e^{-\kappa d}} = e^{2i\theta} \tag{X.108}$$

with the phase

$$\theta = -\arctan \frac{\kappa}{k} \frac{(\kappa + k \cot ka)\, e^{\kappa d} - (\kappa - k \cot ka)\, e^{-\kappa d}}{(\kappa + k \cot ka)\, e^{\kappa d} + (\kappa - k \cot ka)\, e^{-\kappa d}}. \tag{X.109}$$

We check first the case $d = 0$. It follows that

$$\begin{aligned} \theta &= -\arctan \frac{\kappa}{k} \frac{2k \cot ka}{2\kappa} = -\arctan \cot ka \\ &= -\arctan \tan\left(\frac{\pi}{2} + ka\right) = \frac{\pi}{2} + ka. \end{aligned} \tag{X.110}$$

This means that

$$\frac{F'}{G'} = -e^{2ika} \text{ or } \frac{F}{G} = -e^{2ika-2ikb} = -e^{-2ikd} = -1 \tag{X.111}$$

as expected.
Next we examine the case $\kappa d \gg 0$ (whereby we assume $\kappa > 0$). Then it follows from (X.109) due to $e^{-\kappa d} \approx 0$ that:

$$\begin{aligned} &\theta \approx -\arctan \frac{\kappa}{k} \text{ for } \kappa + k \cot ka \neq 0 \\ &\rightarrow \frac{F'}{G'} = \frac{1 + \frac{\kappa}{ik}}{1 - \frac{\kappa}{ik}} \text{ or } \frac{F}{G} = \frac{1 + \frac{\kappa}{ik}}{1 - \frac{\kappa}{ik}} e^{2ikb} \end{aligned} \tag{X.112}$$

$$\begin{aligned} &\theta = \arctan \frac{\kappa}{k} \text{ for } \kappa + k \cot ka = 0 \\ &\rightarrow \frac{F'}{G'} = \frac{1 - \frac{\kappa}{ik}}{1 + \frac{\kappa}{ik}} \text{ or } \frac{F}{G} = \frac{1 - \frac{\kappa}{ik}}{1 + \frac{\kappa}{ik}} e^{2ikb}. \end{aligned} \tag{X.113}$$

So we have a sudden change in the phase at those energies which are determined by $\kappa + k \cot ka = 0$. What is the physical reason?
The equation $\kappa + k \cot ka = 0$ gives the positions of the energy levels in the potential well of length a:

$$V(x) = \begin{cases} 0 & & L < x \\ -V_0 & \text{for} & 0 < x \leq L \\ \infty & & x \leq 0 \end{cases} \tag{X.114}$$

with $V_0 > 0$. Now for the current problem, we do not have bound stable states, i.e. states of infinite lifetime, but nevertheless we find states which have a certain lifetime, called metastable states or resonances. Their energetic positions agree for sufficiently large d approximately with the positions of the bound levels of the potential well (X.114).
Hence, the zeros of the phase (X.109) are crucial for the position of the resonances, i.e.

$$\kappa \sin ka + k \cos ka + (\kappa \sin ka - k \cos ka)\, e^{-2\kappa d} = 0. \tag{X.115}$$

We rewrite this with the abbreviations $e^{-2\kappa d} = \varepsilon$ and $z = ka$. Due to $k^2 = 2mE/\hbar^2$ and $\kappa^2 = 2m\,(V_0 - E)\,/\hbar^2 = 2mV_0/\hbar^2 - k^2$, we arrive with $v^2 = 2ma^2 V_0/\hbar^2$ at

$$\sqrt{v^2 - z^2} \sin z + z \cos z + \left(\sqrt{v^2 - z^2} \sin z - z \cos z\right) \varepsilon = 0;\ 0 \le z \le v. \tag{X.116}$$

We insert $z = z_0 + \varepsilon z_1$ and compare equal powers of ε. In the zeroth approximation, the solutions are determined by

$$\sqrt{v^2 - z_0^2} \sin z_0 + z_0 \cos z_0 = 0;\ 0 \le z_0 \le v. \tag{X.117}$$

For the terms $\sim\varepsilon^1$, we find:

$$\left(\sqrt{v^2 - z_0^2} \cos z_0 + \cos z_0 - \frac{z_0 \sin z_0}{\sqrt{v^2 - z_0^2}} - z_0 \sin z_0\right) z_1 + \sqrt{v^2 - z_0^2} \sin z_0 - z_0 \cos z_0 = 0. \tag{X.118}$$

We replace $\cos z_0$ with the help of (X.117); finally it follows that

$$z_1 = 2z_0 \frac{v^2 - z_0^2}{v^2 \left(1 + \sqrt{v^2 - z_0^2}\right)}. \tag{X.119}$$

These corrections are always positive. This means that the resonances lie at somewhat higher energies than the stable energy levels of the potential (X.114).

9. In this chapter, a transcendental equation of the form

$$\tan kd = -\frac{k}{\kappa} = \ ;\ \kappa = \sqrt{\kappa_V^2 - k^2}\ ;\ k < \kappa_V \tag{X.120}$$

occurs several times. Find an approximate solution for large d.
Solution: We have

$$\tan kd = -\frac{k}{\sqrt{\kappa_V^2 - k^2}} = -\tan\left(\arctan \frac{k}{\sqrt{\kappa_V^2 - k^2}}\right) = -\tan\left(\arcsin \frac{k}{\kappa_V}\right). \tag{X.121}$$

The formulation with arcsin is simpler, because it contains no square root. It follows that:

$$kd + \arcsin \frac{k}{\kappa_V} = n\pi\ ;\ n = 1, 2, \ldots. \tag{X.122}$$

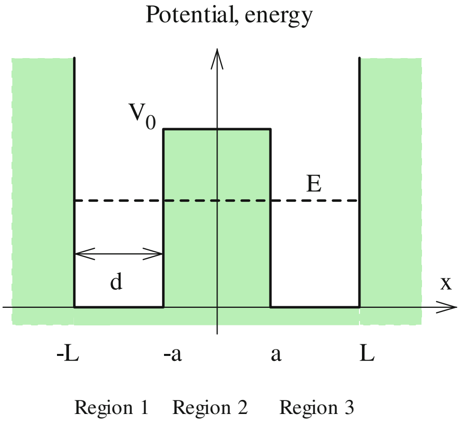


Fig. X.6 Double well potential

Because of $0 \leq \arcsin x \leq \pi/2$ for $x \geq 0$, all solutions k are confined to the intervals $(n - 1/2)\,\pi < kd < n\pi$. Hence, for $d \to \infty$, we have $k \to 0$, and we can use the power series expansion of the Arcus function for small arguments:

$$\arcsin x = x + \frac{x^3}{6} + O\left(x^5\right). \tag{X.123}$$

Neglecting terms of order 5 leads to

$$k\mathrm{d} + \frac{k}{\kappa_V} + \frac{1}{6}\left(\frac{k}{\kappa_V}\right)^3 = n\pi \text{ or } k \cdot \frac{\mathrm{d}\kappa_V + 1}{\kappa_V} = n\pi - \frac{1}{6}\left(\frac{k}{\kappa_V}\right)^3. \tag{X.124}$$

$\left(\frac{k}{\kappa_V}\right)^3$ is a very small term which we approximate using $k = n\pi\kappa_V / \left(\mathrm{d}\kappa_V + 1\right)$. Thus, we obtain

$$\frac{k}{\kappa_V} = \frac{n\pi}{\mathrm{d}\kappa_V + 1}\left[1 - \frac{1}{6}\frac{(n\pi)^2}{(\mathrm{d}\kappa_V + 1)^3}\right]. \tag{X.125}$$

10. Given the double well potential (see Fig. X.6):

$$\begin{array}{ll} \text{region 1: } -L \leq x \leq -a & V = 0 \\ \text{region 2: } -a < x < a & V = V_0 > 0\,. \\ \text{region 3: } a \leq x \leq L & V = 0 \end{array} \tag{X.126}$$

V is infinite for $|x| > L$. We consider only energies E for which $E < V_0$.

(a) Due to the symmetry of the problem ($H\,(x) = H\,(-x)$), there are symmetric and antisymmetric eigenfunctions, sS and aS (cf. Chap. 21). Determine these functions and their eigenvalue equations.
Solution: The *ansatz* for the wavefunction reads

$$\psi_1 = Ae^{ikx} + Be^{-ikx} \ ; \ \psi_2 = Ce^{\kappa x} + De^{-\kappa x} \ ; \ \psi_3 = Fe^{ikx} + Ge^{-ikx}, \tag{X.127}$$

with $k, \kappa > 0$, $0 < E < V_0$ and

$$k^2 = \frac{2m}{\hbar^2}E \ ; \ \kappa_V^2 = \frac{2m}{\hbar^2}V_0 \ ; \ \kappa^2 = \frac{2m}{\hbar^2}V_0 - \frac{2m}{\hbar^2}E = \kappa_V^2 - k^2 > 0. \tag{X.128}$$

The solutions sS and aS have to satisfy the following equation (the upper/lower sign denotes sS/aS)[158]:

$$\psi_1(x) = \pm\psi_3(-x) \ ; \ \psi_2(x) = \pm\psi_2(-x). \tag{X.129}$$

This leads to (region 2 is classically forbidden):

$$\psi_1 = Ae^{ikx} + Be^{-ikx} \ ; \ \psi_2 = Ce^{\kappa x} \pm Ce^{-\kappa x} \ ; \ \psi_3 = \pm Be^{ikx} \pm Ae^{-ikx}. \tag{X.130}$$

The constants B and C are defined by means of the boundary conditions. At $x = -L$, we have

$$\psi_1(-L) = Ae^{-ikL} + Be^{ikL} = 0 \rightarrow B = -Ae^{-2ikL}, \tag{X.131}$$

and at $x = -a$, we have

$$Ae^{-ika} + Be^{ika} = Ce^{-\kappa a} \pm Ce^{\kappa a} \ ; \ ikAe^{-ika} - ikBe^{ika} = \kappa Ce^{-\kappa a} \mp \kappa Ce^{\kappa a}. \tag{X.132}$$

In this way, we obtain for the wavefunction

$$\text{sS:} \ \begin{array}{c} \psi_1 = A_s \sin k\,(x+L) \\ \psi_2 = C_s \cosh \kappa x \\ \psi_3 = -A_s \sin k\,(x-L) \\ C_s = \frac{\sin k(L-a)}{\cosh \kappa a} A_s \end{array} \ ; \text{aS:} \ \begin{array}{c} \psi_1 = A_a \sin k\,(x+L) \\ \psi_2 = C_a \sinh \kappa x \\ \psi_3 = A_a \sin k\,(x-L) \\ C_a = -\frac{\sin k(L-a)}{\sinh \kappa a} A_a \end{array}. \tag{X.133}$$

Therefore, the eigenvalue equations (obtained from (X.132) by division of the two equations) are:

$$\begin{array}{ll} \text{sS:} & \tanh \kappa a \cdot \tan kd = -\frac{k}{\kappa} \\ \text{aS:} & \coth \kappa a \cdot \tan kd = -\frac{k}{\kappa} \end{array} \ ; \ d = L - a, \tag{X.134}$$

which we write as

[158] It goes without saying that one can treat the problem without taking into account the symmetry properties right at the start. In this manner, the symmetry properties emerge by themselves in the course of the computation. However, the calculation is longer and more cumbersome—and the results are identical, of course.

$$\begin{array}{ll} \text{sS:} & kd + \arctan\left(\frac{k}{\kappa}\coth\kappa a\right) = n\pi \\ \text{aS:} & kd + \arctan\left(\frac{k}{\kappa}\tanh\kappa a\right) = n\pi \end{array} \; ; \; n = 1, 2, \ldots \tag{X.135}$$

(b) Show that there is no solution of the eigenvalue equations below a certain threshold value of V_0.
Solution: Due to $0 < \arctan x < \frac{\pi}{2}$ for $x > 0$, from (X.135), it follows immediately that

$$\left(n - \frac{1}{2}\right)\pi < kd < n\pi \; ; \; n = 1, 2, \ldots . \tag{X.136}$$

In particular, for $n = 1$, $\frac{\pi}{2} < kd < \pi$ holds. Thus, because of $0 < k < \kappa_V$, there is no solution of (X.135) for $d\kappa_V < \pi/2$, i.e. when the wells are too narrow and/or the potential V_0 is too low:

$$d^2 V_0 < \frac{\hbar^2\pi^2}{8m} \rightarrow \text{ no solution.} \tag{X.137}$$

(c) Show that the ground state is symmetric.
Solution: Due to $0 < \tanh x < 1$ and $1 < \coth x < \infty$ for all $x > 0$, we can deduce from (X.135) the inequalities

$$\begin{array}{ll} \text{sS:} & \left(n - \frac{1}{2}\right)\pi < kd < n\pi - \arctan\frac{k}{\kappa} \\ \text{aS:} & n\pi - \arctan\frac{k}{\kappa} < kd < n\pi \end{array} \; ; \; n = 1, 2, \ldots . \tag{X.138}$$

We see immediately that the solutions for the symmetric cases are lower than those for the antisymmetric cases of the same order. In other words, the symmetric cases are energetically favorable and the ground state of the system is symmetric.

(d) Solve the eigenvalue equations approximately for the case of a 'thick' barrier, i.e. for very large a.
Solution: The limiting cases of the double well are (i) $\kappa a \rightarrow 0$ and (ii) $\kappa a \rightarrow \infty$, i.e. asymptotically (i) a single potential well of length $2L$ and (ii) two separated potential wells, each of length d. Of course, the position of the energy levels depends on the properties of the barrier. Schematically, this is shown in Fig. X.7.
We concentrate here on the case $\kappa a \gg 1$.
Due to $\tanh x = 1 - 2e^{-2x} + O\left(e^{-4x}\right)$ for $x \rightarrow \infty$, we can approximate (X.134) by (upper/lower sign for sS/aS):

$$kd + \arctan\left(\frac{k}{\kappa}\left(1 \pm 2e^{-2\kappa a}\right)\right) = n\pi \; ; \; n = 1, 2, \ldots . \tag{X.139}$$

For $\kappa a = \infty$, we have $kd + \arctan\frac{k}{\kappa} = n\pi$. Approximate solutions of this equation are given in (X.125); we denote them by $k_{\infty n}$ and write $\kappa_{\infty n} =$

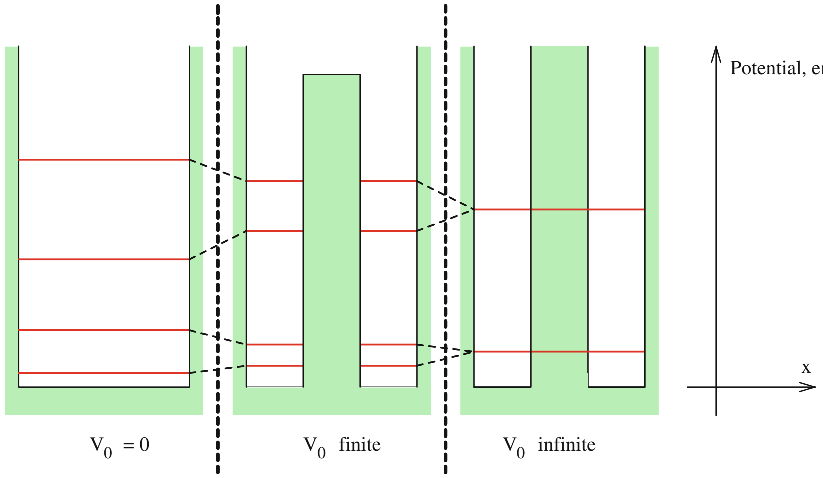


Fig. X.7 Schematic representation of the energy levels in the double well for different barrier heights. Left for $V_0 = 0$, middle for $0 < V_0 < \infty$, right for $V_0 = \infty$. Not to scale

$\sqrt{\kappa_V^2 - k_{\infty n}^2}$. For an approximate solution of (X.139), we insert the *ansatz*

$$k = k_{\infty n} + \hat{k}_n \ ; \ \hat{k}_n = O\left(e^{-2a\kappa_{\infty n}}\right) \ll k_{\infty n} \tag{X.140}$$

and retain only terms $O\left(\hat{k}_n^0\right)$ and $O\left(\hat{k}_n^1\right)$. The result reads (upper/lower sign for sS/aS):

$$\hat{k}_n = \mp k_{\infty n} \frac{\kappa_{\infty n}^2}{\kappa_V^2} \frac{2e^{-2\kappa_\infty a}}{1 + d\kappa_{\infty n}} \text{ or } k = k_{\infty n} \left(1 \mp \frac{\kappa_{\infty n}^2}{\kappa_V^2} \frac{2e^{-2\kappa_{\infty n} a}}{1 + d\kappa_{\infty n}}\right). \tag{X.141}$$

This means that instead of the single energy level $E = \hbar^2 k_{\infty n}^2 / 2m$, we now have a doublet.[159]

As is seen, the determining factor is the exponential function $e^{-2\kappa_{\infty n} a}$, due to which the quantities may react very sensitively to small modifications of the potential. Hence, changing e.g. a or κ_V may have drastic effects.

(e) The initial state is assumed to be a linear combination of the symmetric and the antisymmetric states of the same order (for the sake of simplicity with equal amplitudes, $A_s = A_a = A$). Determine the time behavior of the wavefunction. Calculate the probabilities $P_i(t)$ of finding the object in region i.
Solution: The wavefunction for $t > 0$ is given by:

[159] Thus, in a double well, we observe a splitting of the energy levels into two terms. Correspondingly, in a triple well there is a splitting into three terms, and in an n-fold well into n terms. For large n, this leads to the band structure of solids.

$$\Phi_3(x,t) = -A\sin k_s(x-L)\,e^{i\omega_s t} + A\sin k_a(x-L)\,e^{i\omega_a t}, \tag{X.142}$$

where k_s and k_a are solutions of (X.135); the coefficient may be written as $A = |A|\,e^{i\alpha}$. The probability density for locating the object in region 3 is given by $|\Phi_3(x,t)|^2$, and the corresponding probability by

$$P_3(t) = \int_a^L |\Phi_3(x,t)|^2\,dx. \tag{X.143}$$

Carrying out the integration and introducing the abbreviation

$$\Delta\omega = \omega_a - \omega_s = \frac{\hbar}{2m}\left(k_a^2 - k_s^2\right), \tag{X.144}$$

we obtain finally

$$\begin{aligned} P_3(t)\cdot\frac{2}{d}\frac{1}{|A|^2} &= \left[1-\frac{\sin(2k_s d)}{2k_s d}\right] + \left[1-\frac{\sin(2k_a d)}{2k_a d}\right] \\ &\quad + 2\left[\frac{\sin((k_a+k_s)d)}{(k_a+k_s)\,d} - \frac{\sin((k_a-k_s)d)}{(k_a-k_s)\,d}\right]\cdot\cos(\Delta\omega t). \end{aligned} \tag{X.145}$$

As the calculation shows, $P_1(t)$ has the same time-independent part as $P_3(t)$, while the time-dependent part has the opposite sign from $P_3(t)$. In region 2, the probability P_2 is time independent (but not the probability density):

$$P_2 = |C_s|^2\left(\frac{\sinh 2\kappa_s a}{2\kappa_s} + a\right) + |C_a|^2\left(\frac{\sinh 2\kappa_a a}{2\kappa_a} - a\right) > 0. \tag{X.146}$$

In this way, the total probability $P_1(t) + P_2 + P_3(t)$ is time independent, as it should be.[160]

We see that $P_3(t)$ oscillates with a frequency $\Delta\omega/2\pi$ about the time-independent part of $P_3(t)$. Thus, a part of the position probability swings back and forth between regions 1 and 3. Such behavior is forbidden in classical mechanics, where for $E < V_0$, the two wells are strictly separated even for finite V_0. In quantum mechanics, the two regions 1 and 3 are 'coupled' due to the tunnel effect; this kind of barrier penetration occurs in many different physical situations without having a classical analogue.

(f) In the case of a thick barrier, it holds that $k_a - k_s \ll k_a + k_s$. Calculate up to and including quadratic terms in $k_a - k_s$ the quantities $R_{\max}^{\min} = \min(P_3)/\max(P_3)$ and $\Delta\omega$. Discuss your findings.
Solution: The extrema of $P_3(t)$ are found at $\cos(\Delta\omega t) = \pm 1$. Inserting

---

[160] $|A|^2$ has to be chosen in such a way that the wavefunction is normalized, i.e. so that $P_1(t) + P_2 + P_3(t) = 1$ holds.

(X.141) into (X.145) and expanding in terms of powers of $\hat{k}_n = \frac{k_a - k_s}{2}$, we obtain up to and including terms $O\left(\hat{k}_n^2\right)$[161]:

$$R_{\max}^{\min} = \left(2\hat{k}_n d\right)^2 F\left(2k_{\infty n} d\right) \tag{X.147}$$

with

$$F\left(2k_{\infty n} d\right) = \frac{\frac{1}{6} - \frac{\sin 2k_{\infty n} d}{2k_{\infty n} d} \cdot \left(\frac{1}{(2k_{\infty n} d)^2} - \frac{1}{2}\right) + \frac{\cos 2k_{\infty n} d}{(2k_{\infty n} d)^2}}{2\left(1 - \frac{\sin 2k_{\infty n} d}{2k_{\infty n} d}\right)}. \tag{X.148}$$

Since $F\left(2k_{\infty n} d\right)$ is a bounded and well-behaved function,[162] the behavior of $R_{\max}^{\min}$ is essentially determined by the factor $\left(2\hat{k}_n d\right)^2$. Inserting $\hat{k}_n$ from (X.141), we obtain

$$R_{\max}^{\min} = \left(4\mathrm{d}k_{\infty n} \frac{\kappa_{\infty n}^2}{\kappa_V^2} \frac{e^{-2\kappa_{\infty n} a}}{1 + d\kappa_{\infty n}}\right)^2 \cdot F\left(2k_{\infty n} d\right) \tag{X.149}$$

and

$$\Delta\omega = \frac{4\hbar}{m} k_{\infty n}^2 \frac{\kappa_{\infty n}^2}{\kappa_V^2} \frac{e^{-2\kappa_{\infty n} a}}{1 + d\kappa_{\infty n}}. \tag{X.150}$$

Thus, we have a periodic exchange of probabilities between regions 1 and 3 with the frequency $f = \Delta\omega/2\pi$. If $R_{\max}^{\min}$ is very small, P_3 becomes periodically 'practically' zero, although it never strictly vanishes, due to its definition (X.143). This situation resembles neutrino oscillations (cf. Chap. 8, Vol. 1, exchange of neutrino types) or beats in coupled pendulums (exchange of energy).[163] As mentioned above, the determining factor is the exponential function $e^{-2\kappa_{\infty n} a}$ due to which the quantities may react very sensitively to small changes in the potential.[164]

(g) In the ammonia molecule NH_3, the N atom tunnels back and forth through the plane of the three H atoms. This situation can be modelled by the double well potential with parameters $a = 0.2 \cdot 10^{-10}$ m, $d = 0.3 \cdot 10^{-10}$ m,

[161] Due to the choice $A_s = A_a = A$, there are no terms $O\left(\hat{k}_n^1\right)$.

[162] The function $F(x)$ oscillates with a period $x = 2\pi$; for $x > \pi$, we have $0.02 < F(x) < 0.14$ and $F(x) \to \frac{1}{12}$ for $x \to \infty$. The notation may be simplified somewhat by the use of spherical Bessel functions; see Appendix A, Vol. 2.

[163] Indeed, there may be a difference between the two examples mentioned and the double well, since in the latter, a substantial part of the probability may be contained in region 2.

[164] This sensitivity of the tunnel effect to the potential-barrier properties is also responsible for the enormous range of decay times observed for alpha decay.

$V_0 = 0.255$ eV and $m = 4 \cdot 10^{-27}$ kg (the reduced mass is $\frac{3m_H m_N}{3m_H + m_N}$).[165] Compute numerical values for the ground-state levels, the frequency and $R_{\max}^{\min}$. Discuss your findings.
Solution: Inserting the given data, we obtain

$$\kappa_V = 0.183 \cdot 10^{12} \text{ m}^{-1} ; k_{\infty 1} = 0.880 \cdot 10^{11} \text{ m}^{-1} ; \kappa_{\infty 1} = 0.160 \cdot 10^{12} \text{ m}^{-1}. \tag{X.151}$$

These data give $a\kappa_{\infty 1} = 3.2$, which means that we have a 'thick' barrier and can apply (X.141). It follows (upper/lower signs for sS/aS) that:

$$k_{s,a} = 0.880 \cdot 10^{11} \left(1 \mp 0.433 \cdot 10^{-3}\right) \text{ m}^{-1}. \tag{X.152}$$

For the lowest (unsplit) energy level, we have[166]:

$$E_{\infty 1} = \frac{\hbar^2}{2m} k_{\infty 1}^2 = 0.591 \cdot 10^{-1} \text{ eV} \mathrel{\hat{=}} 0.143 \cdot 10^{14} \text{ Hz}. \tag{X.153}$$

The energy splitting of the lowest level is given by

$$\Delta E_{a,s} = \frac{\hbar^2}{2m} \left(k_a^2 - k_s^2\right) = 0.102 \cdot 10^{-3} \text{ eV} \mathrel{\hat{=}} 0.247 \cdot 10^{11} \text{ Hz} \mathrel{\hat{=}} 0.824 \text{ cm}^{-1}. \tag{X.154}$$

Finally, the ratio $R_{\max}^{\min} = \min(P_3) / \max(P_3)$ has the value

$$R_{\max}^{\min} = 0.253 \cdot 10^{-6}. \tag{X.155}$$

For an intuitive picture, we discuss the findings in terms of coupled pendulums (i.e. probability $\to$ energy).[167] The pendulums oscillate with the frequency $0.143 \cdot 10^{14}$ Hz. The first pendulum pumps energy into the second pendulum until all the energy is in the second pendulum and the first pendulum stops (nearly stops, since we have $0.253 \cdot 10^{-6}$ and not exactly 0). Then the process reverses, is repeated and so on. This continuous 'energy swapping' or beating is a comparatively slow process—we find $\frac{0.143 \cdot 10^{14} \text{ Hz}}{0.247 \cdot 10^{11} \text{ Hz}} = 579$, i.e. a pendulum oscillates several hundred times before the energy is swapped.
For ammonia, this means that the N atom tunnels back and forth through the H_3 plane with a frequency of $0.247 \cdot 10^{11}$ Hz (the 'real' value of this *inversion*

[165] This is a rather rough model, because NH_3 as a three-dimensional molecule has additional degrees of freedom; furthermore, the potential would be better described by a coupling of two harmonic potentials. Hence, one should not take the values for a and d too seriously - they represent simply an order of magnitude. But it is possible to generate a double well potential like (X.126) in the lab. Indeed, there are techniques to construct almost any arbitrary potentials using certain semiconducting materials (keyword heterojunctions or heterostructures).

[166] For the conversion of the energy units, see e.g. Appendix A, Vol. 1.

[167] We can do this because the contribution of the classically-forbidden region 2 for the given data is $P_2 \approx 0.02$, which we can neglect in this context.

frequency is $0.2387 \cdot 10^{11}$ Hz or $0.8\ \text{cm}^{-1}$).[168] The probability of finding the N on one side of the H_3 plane varies periodically from nearly 1 to nearly 0 and back again during this process.[169]

11. For an illustration of the method of stationary phase, consider the (unnormalized) wavefunction

$$\psi(x,t) = \int_{-\infty}^{\infty} |A(k)|\, e^{i\varphi(k)} e^{i(kx-\omega t)} \mathrm{d}k \tag{X.156}$$

with

$$\omega = ck;\ \ \varphi(k) = -x_0 k \tag{X.157}$$

and

$$|A(k)| = \begin{cases} \kappa^2 - (k-K)^2 & \text{for} \quad 0 < K - \kappa \leq k \leq K + \kappa \\ 0 & \qquad\qquad \text{otherwise} \end{cases}. \tag{X.158}$$

The constants κ, K and x_0 are positive. Calculate explicitly $\psi(x,t)$ and discuss its properties. What is the physical significance of x_0?
Solution: We have to evaluate

$$\psi(x,t) = \int_{K-\kappa}^{K+\kappa} \left[\kappa^2 - (k-K)^2\right] e^{ik(x-ct-x_0)} \mathrm{d}k. \tag{X.159}$$

To simplify the notation, we introduce the abbreviation $\delta = x - ct - x_0$ and substitute $z = k - K$ in the integral. This leads to

$$\psi(x,t) = e^{iK\delta} \int_{-\kappa}^{\kappa} \left(\kappa^2 - z^2\right) e^{iz\delta} \mathrm{d}z. \tag{X.160}$$

---

[168] We mention two applications: (i) This frequency is also used to identify ammonia in interstellar space (radio astronomy). (ii) Since NH_3 is polar, there is an oscillating dipole moment associated with the tunnelling through the H_3 plane, which fact is used in the ammonia *maser* (acronym for *Microwave Amplification by Stimulated Emission of Radiation*). The maser action in fact takes place between the two lowest levels considered above. Using external fields, one generates a population inversion with respect to these levels, followed by stimulated emission (with the frequency $0.247 \cdot 10^{11}$ Hz, i.e. in the microwave range). The extension of this concept to the realm of light (and, of course, to other materials) leads to the *laser* (*Light Amplification by Stimulated Emission of Radiation*).

[169] Structurally similar molecules demonstrate the sensitive dependence of tunnelling on the barrier potential in an exemplary manner. Thus, NH_3, PH_3 and AsH_3, with barrier heights of 0.25 eV, 0.75 eV and 1.39 eV, have inversion frequencies of $0.24 \cdot 10^{11}$ Hz, $0.14 \cdot 10^{6}$ Hz and $0.16 \cdot 10^{-7}$ Hz (i.e. 0.5 per year).

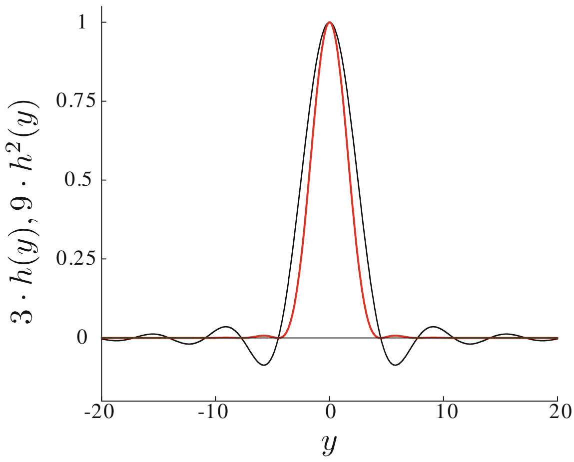


Fig. X.8 The functions $h(y)$ (*black*), (X.162) and $h^2(y)$ (*red*)

The integral may be evaluated by hand, by using software such as *Maple* or *mathematica*, or by an online integrator (e.g. http://integrals.wolfram.com/). The result can be brought into the form:

$$\psi(x,t) = 4e^{iK\delta}\kappa^3 \cdot \frac{\sin\kappa\delta - \kappa\delta\cos\kappa\delta}{(\kappa\delta)^3}. \tag{X.161}$$

As we see, $\psi(x,t)$ depends only on $\delta = x - ct - x_0$. Hence, the wavefunction moves along the x axis without dispersion, i.e. without changing its shape in the course of time, $\psi(x,t) = \psi(x-t)$. The function

$$h(y) = \frac{\sin y - y\cos y}{y^3} \tag{X.162}$$

on the right-hand side of (X.161) determines the behavior of $\psi(x,t)$. It is shown in Fig. X.8.[170] The function has a maximum at $y = 0$ with $h(0) = \frac{1}{3}$, and a halfwidth of $\Delta y \approx 5$ (i.e. $h(\pm 2.498) \approx \frac{1}{6}$); the first zero lies at $y \approx 4.493$. With $\delta = x - ct - x_0$, this means that $\psi(x,t)$ has a pronounced maximum at $x - ct - x_0 = 0$. Thus, x_0 is the position of the maximum at the time t=0. The halfwidth of $\psi(x,t)$ is given by $\Delta x \approx \frac{5}{\kappa}$ and, correspondingly, the halfwidth of $|\psi(x,t)|^2$ is given by $\Delta x \approx \frac{3.6}{\kappa}$. In other words, the integral (X.156) yields essential contributions only in the neighborhood of the stationary phase $\frac{\mathrm{d}}{\mathrm{d}k}(kx - \omega t + \varphi(k)) = 0$.

[170] Note, by the way, that $h(y) = \frac{j_1(y)}{y}$, where $j_1(y)$ is a spherical Bessel function.

X.2 Exercises, Chap. 16

1. For which K, N, M are the spherical harmonics (in spherical coordinates)

$$f(\vartheta, \varphi) = \cos^K \vartheta \cdot \sin^M \vartheta \cdot e^{iN\varphi} \tag{X.163}$$

eigenfunctions of $\mathbf{l}^2$?

2. Write out the spherical harmonics for $l = 1$ using Cartesian coordinates, x, y, z.
Solution: With $x = r \sin\vartheta \cos\varphi$, $y = r \sin\vartheta \sin\varphi$, $z = r \cos\vartheta$, it follows that

$$Y_1^0(\vartheta, \varphi) = \sqrt{\frac{3}{4\pi}} \frac{z}{r}; \; Y_1^{\pm 1}(\vartheta, \varphi) = \mp\sqrt{\frac{3}{8\pi}} \frac{x \pm iy}{r}. \tag{X.164}$$

3. Show that:

$$\mathbf{l} \cdot \hat{\mathbf{r}} = \hat{\mathbf{r}} \cdot \mathbf{l} = 0 \tag{X.165}$$

4. Show that the components of $\mathbf{l}$ are Hermitian.
Solution: We have

$$l_x = \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) = yp_z - zp_y. \tag{X.166}$$

On the right side, we have products of two commuting Hermitian operators. Hence, l_x is Hermitian. Analogously for the two other components of $\mathbf{l}$.

5. Show that for the orbital angular momentum, it holds that

$$\left[l_x, l_y\right] = i\hbar l_z; \; \left[l_y, l_z\right] = i\hbar l_x; \; \left[l_z, l_x\right] = i\hbar l_y. \tag{X.167}$$

Solution: We first consider $\left[l_x, l_y\right]$. We have:

$$\begin{aligned}
[l_x, l_y] &= \left(\frac{\hbar}{i}\right)^2 \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) - \left(\frac{\hbar}{i}\right)^2 \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \\
&= \left(\frac{\hbar}{i}\right)^2 \left(y\frac{\partial}{\partial z}z\frac{\partial}{\partial x} - y\frac{\partial}{\partial z}x\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}z\frac{\partial}{\partial x} + z\frac{\partial}{\partial y}x\frac{\partial}{\partial z}\right) \\
&\quad - \left(\frac{\hbar}{i}\right)^2 \left(z\frac{\partial}{\partial x}y\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}z\frac{\partial}{\partial y} - x\frac{\partial}{\partial z}y\frac{\partial}{\partial z} + x\frac{\partial}{\partial z}z\frac{\partial}{\partial y}\right) \\
&= \left(\frac{\hbar}{i}\right)^2 \left(y\frac{\partial}{\partial z}z\frac{\partial}{\partial x} + z\frac{\partial}{\partial y}x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}y\frac{\partial}{\partial z} - x\frac{\partial}{\partial z}z\frac{\partial}{\partial y}\right) \\
&= \left(\frac{\hbar}{i}\right)^2 \left(y\frac{\partial}{\partial z}\frac{\partial}{\partial x} + yz\frac{\partial}{\partial z}z\frac{\partial}{\partial x} + zx\frac{\partial}{\partial y}\frac{\partial}{\partial z} - zy\frac{\partial}{\partial x}\frac{\partial}{\partial z} - x\frac{\partial}{\partial z}\frac{\partial}{\partial y} - xz\frac{\partial}{\partial z}z\frac{\partial}{\partial y}\right) \\
&= \left(\frac{\hbar}{i}\right)^2 \left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right) = -\frac{\hbar}{i}l_z = i\hbar l_z.
\end{aligned} \tag{X.168}$$

The other two relations follow by cyclic commutation.
An alternative derivation uses the commutators $[x, p_x] = i\hbar$, $\left[x, p_y\right] = \left[x, p_z\right] = 0$; correspondingly for y, z. We have:

$$\begin{aligned}\left[l_x, l_y\right] &= \left[yp_z - zp_y, zp_x - xp_z\right] \\ &= \left[yp_z, zp_x\right] - \left[yp_z, xp_z\right] - \left[zp_y, zp_x\right] + \left[zp_y, xp_z\right].\end{aligned} \tag{X.169}$$

The second and the third commutators vanish due to $\left[yp_z, xp_z\right] = yx\left[p_z, p_z\right] = 0$ and $\left[zp_y, zp_x\right] = z^2\left[p_y, p_x\right] = 0$. Then it follows that:

$$\left[l_x, l_y\right] = y\left[p_z, zp_x\right] + x\left[zp_y, p_z\right] = y\left[p_z, zp_x\right] - x\left[p_z, zp_y\right]. \tag{X.170}$$

With $[A, BC] = B[A, C] + [A, B]C$ (see next exercise), we obtain

$$\left[l_x, l_y\right] = yz\left[p_z, p_x\right] + y\left[p_z, z\right]p_x - xz\left[p_z, p_y\right] - x\left[p_z, z\right]p_y. \tag{X.171}$$

The first and the third commutator vanish; due to $\left[p_z, z\right] = -i\hbar$, it follows that:

$$\left[l_x, l_y\right] = -i\hbar yp_x + i\hbar xp_y = i\hbar\left(xp_y - yp_x\right) = i\hbar l_z. \tag{X.172}$$

6. Show that $[A, BC] = B[A, C] + [A, B]C$ holds. Using this identity and the commutators $\left[l_x, l_y\right] = i\hbar l_z$ plus cyclic permutations, prove that $\left[l_x, \mathbf{l}^2\right] = 0$.
7. Show that:

$$\left[\mathbf{J}^2, J_\pm\right] = 0. \tag{X.173}$$

8. We have seen in the text that

$$J_\pm\left|j, m\right\rangle = c^\pm_{j,m}\left|j, m \pm 1\right\rangle. \tag{X.174}$$

Using

$$\begin{aligned}J_+J_-\left|j, m\right\rangle &= \hbar^2\left[j(j+1) - m(m-1)\right]\left|j, m\right\rangle \\ J_-J_+\left|j, m\right\rangle &= \hbar^2\left[j(j+1) - m(m+1)\right]\left|j, m\right\rangle,\end{aligned} \tag{X.175}$$

show that for the coefficients $c^\pm_{j,m}$,

$$c^\pm_{j,m} = \hbar\sqrt{j(j+1) - m(m \pm 1)} \tag{X.176}$$

holds.
Solution: We consider the first equation of (X.175). It follows that

$$J_+J_-\left|j, m\right\rangle = J_+c^-_{j,m}\left|j, m-1\right\rangle = c^+_{j,m-1}c^-_{j,m}\left|j, m\right\rangle \tag{X.177}$$

and thus,

$$c^{+}_{j,m-1}c^{-}_{j,m} = \hbar^2 \left[j(j+1) - m(m-1) \right]. \tag{X.178}$$

(The second equation yields the same result for the index $m+1$ instead of m). We use the *ansatz*

$$c^{+}_{j,m} = \hbar\sqrt{j\,(j+1) - d^{+}_{j,m}};\ c^{-}_{j,m} = \hbar\sqrt{j\,(j+1) - d^{-}_{j,m}} \tag{X.179}$$

and obtain

$$\sqrt{j\,(j+1) - d^{+}_{j,m-1}}\sqrt{j\,(j+1) - d^{-}_{j,m}} = j(j+1) - m(m-1). \tag{X.180}$$

Squaring and multiplying yields

$$\begin{aligned} &j\,(j+1)\,d^{-}_{j,m} + j\,(j+1)\,d^{+}_{j,m-1} - d^{+}_{j,m-1}d^{-}_{j,m} \\ &= 2j(j+1)m(m-1) - m^2(m-1)^2. \end{aligned} \tag{X.181}$$

Comparing equal powers of j leads to

$$d^{+}_{j,m-1} + d^{-}_{j,m} = 2m(m-1);\ d^{+}_{j,m-1}d^{-}_{j,m} = m^2(m-1)^2 \tag{X.182}$$

and this yields

$$d^{+}_{j,m-1} = d^{-}_{j,m} = m(m-1) \tag{X.183}$$

or

$$c^{\pm}_{j,m} = \hbar\sqrt{j\,(j+1) - m\,(m \pm 1)}. \tag{X.184}$$

9. Given the Pauli matrices σ_k,
   (a) Show (once more) that

   $$\left[\sigma_i, \sigma_j\right] = 2i\varepsilon_{ijk}\sigma_k;\ \left\{\sigma_i, \sigma_j\right\} = 2\delta_{ij};\ \sigma_i^2 = 1;\ \sigma_i\sigma_j = i\varepsilon_{ijk}\sigma_k; \tag{X.185}$$

   (b) Prove that

   $$(\boldsymbol{\sigma}\mathbf{A})\,(\boldsymbol{\sigma}\mathbf{B}) = \mathbf{AB} + i\boldsymbol{\sigma}\,(\mathbf{A} \times \mathbf{B}) \tag{X.186}$$

   where $\boldsymbol{\sigma}$ is the vector $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ and $\mathbf{A}, \mathbf{B}$ are three-dimensional vectors;
   (c) Show that every 2×2 matrix can be expressed as a linear combination of the three Pauli matrices and the unit matrix.

10. Given the orbital angular momentum operator $\mathbf{l}$ and the spin operator $\mathbf{s}$, show that $\left[l_z, \mathbf{s}\cdot\mathbf{l}\right] \neq 0;\ \left[s_z, \mathbf{s}\cdot\mathbf{l}\right] \neq 0;\ \left[l_z + s_z, \mathbf{s}\cdot\mathbf{l}\right] = 0$.
    Solution: We have

$$\begin{aligned}\left[l_z, \mathbf{s}\cdot\mathbf{l}\right] &= s_x\left[l_z, l_x\right] + s_y\left[l_z, l_y\right] = i\hbar\left(s_x l_y - s_y l_x\right)\\ \left[s_z, \mathbf{s}\cdot\mathbf{l}\right] &= l_x\left[s_z, s_x\right] + l_y\left[s_z, s_y\right] = i\hbar\left(l_x s_y - l_y s_x\right)\end{aligned} \tag{X.187}$$

and from this, $\left[l_z + s_z, \mathbf{s}\cdot\mathbf{l}\right] = 0$.

11. The ladder operators for a generalized angular momentum are given as $J_\pm = J_x \pm iJ_y$.

    (a) Show that $\left[J_z, J_+\right] = \hbar J_+$, $\left[J_z, J_-\right] = -\hbar J_-$, $\left[J_+, J_-\right] = 2\hbar J_z$, as well as $\mathbf{J}^2 = \frac{1}{2}\left(J_+J_- + J_-J+\right) + J_z^2$.
    (b) Show that it follows from the last equation that:

$$J_+J_- = \mathbf{J}^2 - J_z\left(J_z - \hbar\right);\; J_-J_+ = \mathbf{J}^2 - J_z\left(J_z + \hbar\right) \tag{X.188}$$

    and hence

$$\begin{aligned}J_+J_-\left|j,m\right\rangle &= \hbar^2\left[j(j+1) - m(m-1)\right]\left|j,m\right\rangle\\ J_-J_+\left|j,m\right\rangle &= \hbar^2\left[j(j+1) - m(m+1)\right]\left|j,m\right\rangle.\end{aligned} \tag{X.189}$$

    (c) Show that from the last two equations, it follows that:

$$\begin{aligned}j(j+1) - m(m-1) &= (j-m)(j+m+1) \geq 0\\ j(j+1) - m(m+1) &= (j+m)(j-m+1) \geq 0\end{aligned} \tag{X.190}$$

    and hence

$$-j \leq m \leq j. \tag{X.191}$$

12. What is the matrix representation of the orbital angular momentum for $l = 1$?
    Solution: We start from

$$J_\pm\left|j,m\right\rangle = \hbar\sqrt{j\left(j+1\right) - m\left(m\pm 1\right)}\left|j,m\pm 1\right\rangle \tag{X.192}$$

    and obtain as a first step:

$$\begin{aligned}\left\langle 1,m'\right| L_\pm\left|1,m\right\rangle &= \hbar\sqrt{2 - m\left(m\pm 1\right)}\delta_{m',m\pm 1}\\ \left\langle 1,m'\right| L_z\left|1,m\right\rangle &= \hbar m\delta_{m',m}.\end{aligned} \tag{X.193}$$

    It follows that

$$\begin{aligned}&\left\langle 1,1\right| L_+\left|1,0\right\rangle = \hbar\sqrt{2};\;\; \left\langle 1,0\right| L_+\left|1,-1\right\rangle = \hbar\sqrt{2}\\ &\left\langle 1,0\right| L_-\left|1,1\right\rangle = \hbar\sqrt{2};\;\; \left\langle 1,-1\right| L_-\left|1,0\right\rangle = \hbar\sqrt{2}\\ &\left\langle 1,1\right| L_z\left|1,1\right\rangle = \hbar;\;\; \left\langle 1,-1\right| L_z\left|1,-1\right\rangle = -\hbar.\end{aligned} \tag{X.194}$$

    All other matrix elements vanish.

We expand (X.193) into

$$\begin{aligned}&\sum_{m,m'} \left|1,m'\right\rangle\left\langle 1,m'\right| L_\pm \left|1,m\right\rangle\left\langle 1,m\right| \\ &\qquad = \sum_{m,m'} \hbar\sqrt{2-m\left(m\pm 1\right)}\delta_{m',m\pm 1}\left|1,m'\right\rangle\left\langle 1,m\right| \\ &\qquad \sum_{m,m'} \left|1,m'\right\rangle\left\langle 1,m'\right| L_z \left|1,m\right\rangle\left\langle 1,m\right| = \sum_{m,m'} \hbar m\delta_{m',m}\left|1,m'\right\rangle\left\langle 1,m\right|.\end{aligned} \tag{X.195}$$

Since $\{|1,m\rangle\}$ is a CONS, it follows that

$$\begin{aligned}L_\pm &= \hbar\sum_m \sqrt{2-m\left(m\pm 1\right)}\left|1,m\pm 1\right\rangle\left\langle 1,m\right| \\ L_z &= \hbar\sum_m m\left|1,m\right\rangle\left\langle 1,m\right|,\end{aligned} \tag{X.196}$$

or, explicitly,

$$\begin{aligned}L_+ &= \hbar\sqrt{2}\left|1,1\right\rangle\left\langle 1,0\right| + \hbar\sqrt{2}\left|1,0\right\rangle\left\langle 1,-1\right| \\ L_- &= \hbar\sqrt{2}\left|1,0\right\rangle\left\langle 1,1\right| + \hbar\sqrt{2}\left|1,-1\right\rangle\left\langle 1,0\right| = L_+^\dagger \\ L_z &= \hbar\left|1,1\right\rangle\left\langle 1,1\right| - \hbar\left|1,-1\right\rangle\left\langle 1,-1\right|.\end{aligned} \tag{X.197}$$

Evidently, the explicit matrix form of the angular momentum operators depends on the representation of the basis vectors. If we choose e.g.

$$\left|1,1\right\rangle \cong \begin{pmatrix}1\\0\\0\end{pmatrix};\ \left|1,0\right\rangle \cong \begin{pmatrix}0\\1\\0\end{pmatrix};\ \left|1,-1\right\rangle \cong \begin{pmatrix}0\\0\\1\end{pmatrix} \tag{X.198}$$

then it follows that

$$L_+ \cong \hbar\sqrt{2}\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix};\ L_- \cong \hbar\sqrt{2}\begin{pmatrix}0&0&0\\1&0&0\\0&1&0\end{pmatrix};\ L_z \cong \hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix} \tag{X.199}$$

and thus,

$$L_x \cong \frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix};\ L_y \cong \frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&-i&0\\i&0&-i\\0&i&0\end{pmatrix};\ L_z \cong \hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}. \tag{X.200}$$

Another choice of the basis vectors is

$$|1,1\rangle \cong \pm\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\\ 0\end{pmatrix};\;\; |1,0\rangle \cong \begin{pmatrix}0\\ 0\\ 1\end{pmatrix};\;\; |1,-1\rangle \cong \mp\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\\ 0\end{pmatrix}. \quad \text{(X.201)}$$

It leads to

$$L_+ \cong \pm\hbar\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & i\\ -1 & -i & 0\end{pmatrix};\; L_- \cong \pm\hbar\begin{pmatrix}0 & 0 & -1\\ 0 & 0 & i\\ 1 & -i & 0\end{pmatrix};\; L_z \cong \hbar\begin{pmatrix}0 & -i & 0\\ i & 0 & 0\\ 0 & 0 & 0\end{pmatrix} \quad \text{(X.202)}$$

or

$$L_x \cong \pm\hbar\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & i\\ 0 & -i & 0\end{pmatrix};\; L_y \cong \pm\hbar\begin{pmatrix}0 & 0 & -i\\ 0 & 0 & 0\\ i & 0 & 0\end{pmatrix};\; L_z \cong \hbar\begin{pmatrix}0 & -i & 0\\ i & 0 & 0\\ 0 & 0 & 0\end{pmatrix}. \quad \text{(X.203)}$$

13. Consider the orbital angular momentum $l = 1$. Express the operator $e^{-i\alpha L_z/\hbar}$ as sum over dyadic products (representation-free). Specify for the bases (X.198) and (X.201).
Solution: We have initially

$$e^{-i\alpha L_z/\hbar} = \sum \frac{(-i\alpha)^n}{n!}\left(\frac{L_z}{\hbar}\right)^n. \quad \text{(X.204)}$$

With $L_z = \hbar\left[|1,1\rangle\langle 1,1| - |1,-1\rangle\langle 1,-1|\right]$ (see (X.197)), it follows that

$$\begin{aligned}\left(\frac{L_z}{\hbar}\right)^2 &= \left[|1,1\rangle\langle 1,1| - |1,-1\rangle\langle 1,-1|\right]^2 = |1,1\rangle\langle 1,1| + |1,-1\rangle\langle 1,-1|\\ \left(\frac{L_z}{\hbar}\right)^3 &= \left[|1,1\rangle\langle 1,1| - |1,-1\rangle\langle 1,-1|\right]\left[|1,1\rangle\langle 1,1| + |1,-1\rangle\langle 1,-1|\right] = \frac{L_z}{\hbar}.\end{aligned} \quad \text{(X.205)}$$

Due to

$$e^{-i\alpha L_z/\hbar} = 1 + \sum_{n=1} \frac{(-i\alpha)^{2n}}{(2n)!}\left(\frac{L_z}{\hbar}\right)^{2n} + \sum_{n=0} \frac{(-i\alpha)^{2n+1}}{(2n+1)!}\left(\frac{L_z}{\hbar}\right)^{2n+1} \quad \text{(X.206)}$$

we obtain

$$e^{-i\alpha L_z/\hbar} = 1 + \left(\frac{L_z}{\hbar}\right)^2 \sum_{n=1} (-1)^n \frac{\alpha^{2n}}{(2n)!} - i\left(\frac{L_z}{\hbar}\right) \sum_{n=0} (-1)^n \frac{\alpha^{2n+1}}{(2n+1)!}. \quad \text{(X.207)}$$

This yields

$$e^{-i\alpha L_z/\hbar} = 1 + \left(\frac{L_z}{\hbar}\right)^2 (\cos\alpha - 1) - i\left(\frac{L_z}{\hbar}\right)\sin\alpha$$
$$= |1,1\rangle\langle 1,1| e^{-i\alpha} + |1,0\rangle\langle 1,0| + |1,-1\rangle\langle 1,-1| e^{i\alpha}. \quad \text{(X.208)}$$

The choice of basis (X.198) yields the representation

$$e^{-i\alpha L_z/\hbar} \underset{\text{(W.162)}}{\cong} \begin{pmatrix} e^{-i\alpha} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\alpha} \end{pmatrix}, \quad \text{(X.209)}$$

and the choice of basis (X.201) yields the representation

$$e^{-i\alpha L_z/\hbar} \underset{\text{(W.165)}}{\cong} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad \text{(X.210)}$$

14. Calculate the term

$$e^{-i\frac{\gamma\hat{\mathbf{a}}L}{\hbar}} = e^{-i\gamma\hat{\mathbf{a}}l} \quad \text{(X.211)}$$

for the orbital angular momentum $l = 1$ and the basis (16.73).[171] $\hat{\mathbf{a}}$ is the rotation axis (a unit vector), γ the rotation angle. For reasons of economy, use the 'simplified' angular momentum $\mathbf{l} = \mathbf{L}/\hbar$, i.e. the theoretical units system.

(a) Express the rotations about the x-, y- and z-axis as matrices.
Solution: The matrix representation of the angular momentum is

$$l_x \cong \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & -i & 0 \end{pmatrix};\; l_y \cong \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix};\; l_z \cong \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad \text{(X.212)}$$

Due to

$$l_x^2 \cong \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix};\; l_y^2 \cong \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix};\; l_z^2 \cong \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{(X.213)}$$

we have the relations

$$l_x^3 = l_x;\; l_y^3 = l_y\;;\; l_z^3 = l_z. \quad \text{(X.214)}$$

For e^{-ibl_k}, it follows due to $l_i^4 = l_i^2$ etc. that:

[171] Of course, all the calculations may also be performed representation-free.

$$e^{-ibl_k} = \sum_{n=0} \frac{(-ib)^n}{n!} l_i^n = 1 + l_k^2 \left[\sum_{n=0} \frac{(-1)^n b^{2n}}{(2n)!} - 1 \right] - i \sum_{n=0} \frac{(-1)^n b^{2n+1}}{(2n+1)!} l_k \tag{X.215}$$

and from this:

$$e^{-ibl_k} = 1 + [\cos b - 1] \cdot l_k^2 - i \sin b \cdot l_k. \tag{X.216}$$

(b) Express the rotations about an axis $\hat{a}$ with rotation angle γ as matrices (the angles in spherical coordinates are θ and φ; see Fig. X.9).
Solution: With the representation (X.212), we obtain

$$\hat{\mathbf{a}}\mathbf{l} = i \begin{pmatrix} 0 & -a_z & -a_y \\ a_z & 0 & a_x \\ a_y & -a_x & 0 \end{pmatrix}. \tag{X.217}$$

Since $\hat{\mathbf{a}}$ is an unit vector, we have $a_x^2 + a_y^2 + a_z^2 = 1$. For $\left(\hat{\mathbf{a}}l\right)^2$, the calculation gives (check yourself):

$$\left(\hat{\mathbf{a}}\mathbf{l}\right)^2 = \begin{pmatrix} 1 - a_x^2 & -a_x a_y & a_x a_z \\ -a_x a_y & 1 - a_y^2 & a_y a_z \\ a_x a_z & a_y a_z & 1 - a_z^2 \end{pmatrix} \tag{X.218}$$

and it follows that

$$\left(\hat{\mathbf{a}}\mathbf{l}\right)^3 = \hat{\mathbf{a}}\mathbf{l} \tag{X.219}$$

This yields the following relation, corresponding to (X.216):

$$e^{-i\gamma\hat{\mathbf{a}}\mathbf{l}} = \sum_{n=0} \frac{(-i\gamma)^n}{n!} \left(\hat{\mathbf{a}}\mathbf{l}\right)^n = 1 + [\cos\gamma - 1] \cdot \left(\hat{\mathbf{a}}\mathbf{l}\right)^2 - i \sin\gamma \cdot \left(\hat{\mathbf{a}}\mathbf{l}\right). \tag{X.220}$$

In matrix representation, we obtain

$$e^{-i\gamma\hat{\mathbf{a}}\mathbf{l}} \cong 1 + [\cos\gamma - 1] \begin{pmatrix} 1 - a_x^2 & -a_x a_y & a_x a_z \\ -a_x a_y & 1 - a_y^2 & a_y a_z \\ a_x a_z & a_y a_z & 1 - a_z^2 \end{pmatrix} + \sin\gamma \begin{pmatrix} 0 & -a_z & -a_y \\ a_z & 0 & a_x \\ a_y & -a_x & 0 \end{pmatrix}. \tag{X.221}$$

To cast this expression in a more familiar form, we rewrite:

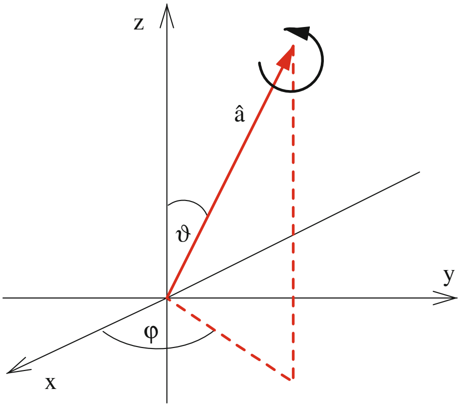


Fig. X.9 Rotation about an axis $\hat{a}$

$$e^{-i\gamma\hat{\mathbf{a}}\mathbf{l}} \cong \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} + (1-\cos\gamma)\begin{pmatrix} a_x^2 & a_x a_y & -a_x a_z \\ a_x a_y & a_y^2 & -a_y a_z \\ -a_x a_z & -a_y a_z & -\left(a_x^2 + a_y^2\right) \end{pmatrix}$$
$$+ \sin\gamma \begin{pmatrix} 0 & 1-a_z & -a_y \\ a_z - 1 & 0 & a_x \\ a_y & -a_x & 0 \end{pmatrix}. \tag{X.222}$$

We recognize in the first matrix the well-known formulation for a two-dimensional rotation about the z-axis.
Finally, we can insert into (X.221) the components of the rotation axis explicitly, i.e. (see Fig. X.9):

$$a_x = \sin\theta\cos\varphi;\;\; a_y = \sin\theta\sin\varphi;\;\; a_z = \cos\theta. \tag{X.223}$$

We do not do this in the general case, as fairly extensive expressions[172] result; instead, we consider only two special cases.
First, we insert $\theta = 0$ or $a_z = 1$; then follows the familiar representation:

$$e^{-i\gamma\hat{\mathbf{a}}\mathbf{l}} \underset{a_z=1}{\cong} \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{X.224}$$

The next case is $\varphi = 0$ or $a_y = 0$; it follows that

$$e^{-i\gamma\hat{\mathbf{a}}\mathbf{l}} \underset{a_y=0}{\cong} \begin{pmatrix} 1 + (\cos\gamma - 1)\cos^2\theta & -\sin\gamma\cos\theta & (\cos\gamma - 1)\sin\theta\cos\theta \\ \sin\gamma\cos\theta & \cos\gamma & \sin\gamma\sin\theta \\ (\cos\gamma - 1)\sin\theta\cos\theta & -\sin\gamma\sin\theta & 1 + (\cos\gamma - 1)\sin^2\theta \end{pmatrix}. \tag{X.225}$$

---

[172] It is remarkable how much more complicated it is to describe rotations in 3 than in 2 dimensions.

If we choose here $\theta = \frac{\pi}{2}$ or $a_x = 1$ (i.e. $a_y = a_z = 0$), we obtain

$$e^{-i\gamma\hat{\mathbf{a}}\mathbf{l}} \underset{a_x=1}{\cong} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & \sin\gamma \\ 0 & -\sin\gamma & \cos\gamma \end{pmatrix}. \tag{X.226}$$

Finally, we want to remark that we can also describe the rotation $e^{-i\gamma\hat{\mathbf{a}}\mathbf{l}}$ as follows: We first rotate the axis by the angle $-\varphi$ about the z-axis (then $\hat{a}$ lies around the x -axis), then by $-\theta$ about the y-axis ($\hat{a}$ now coincides with the z-axis). Now we can perform the rotation about the z-axis by γ and then reverse the rotations by θ and φ. This yields

$$e^{-i\gamma\hat{\mathbf{a}}\mathbf{l}} = e^{-i\varphi l_z} e^{-i\theta l_y} e^{-i\gamma l_z} e^{i\theta l_y} e^{i\varphi l_z}. \tag{X.227}$$

One can evaluate this expression with (X.216) and again obtains (X.221).

X.3 Exercises, Chap. 17

1. Derive (17.14) from (17.11).
2. Show that

$$u_{E;l}(r) \underset{r\to\infty}{\sim} r^{\alpha} \text{ with } \alpha < -\frac{1}{2} \tag{X.228}$$

must hold.
Solution: In order that the wavefunction is square integrable, and due to $\mathrm{d}V = r^2 \mathrm{d}r \sin\vartheta \, d\vartheta \, d\varphi$, it must hold that

$$|\psi|^2 r^2 \sim R_{nl}^2 r^2 \sim r^b \text{ with } b < -1. \tag{X.229}$$

Because of $u_{E;l} = rR_{E;l}$, the proposition follows directly.
3. Hydrogen atom: the probability density of the electron in a volume element $\mathrm{d}^3r = r^2\mathrm{d}r\mathrm{d}\Omega$ around the point (r, ϑ, φ) is given by

$$\mathrm{d}^3 w(r, \vartheta, \varphi) = |R_{nl}(r)|^2 \left|Y_l^m(\vartheta, \varphi)\right|^2 r^2 \mathrm{d}r\mathrm{d}\Omega = |u_{nl}(r)|^2 \left|Y_l^m(\vartheta, \varphi)\right|^2 \mathrm{d}r\mathrm{d}\Omega. \tag{X.230}$$

Find graphical representations, as illustrative as possible, of the probability densities for the various orbitals with $n = 1$ and $n = 2$.

X.4 Exercises, Chap. 18

1. Show explicitly that the eigenvalues of $\hat{n}$ are positive.
Solution: Since $\hat{n}$ is Hermitian, the eigenvalues are real. In addition, we have:

$$\begin{aligned} \langle\nu|\,\hat{n}\,|\nu\rangle &= \nu\,\langle\nu\,|\nu\rangle = \nu \\ \langle\nu|\,\hat{n}\,|\nu\rangle &= \langle\nu|\,a^{\dagger}a\,|\nu\rangle = \|a\,|\nu\rangle\|^2 \geq 0 \end{aligned} \tag{X.231}$$

and therefore $\nu \geq 0$.

2. Show that

$$a^l\,|\nu\rangle = \sqrt{(\nu-l)!\binom{\nu}{l}}\,|\nu-l\rangle \text{ and } a^{\dagger k}\,|\nu\rangle = \sqrt{\nu!\binom{\nu+k}{k}}\,|\nu+k\rangle\,. \tag{X.232}$$

Solution: We start from:

$$a\,|\nu\rangle = \sqrt{\nu}\,|\nu-1\rangle\,;\ a^{\dagger}\,|\nu\rangle = \sqrt{\nu+1}\,|\nu+1\rangle$$

and thus

$$\begin{aligned} a^l\,|\nu\rangle &= \sqrt{\nu\,(\nu-1)\ldots(\nu-l+1)}\,|\nu-l\rangle = \sqrt{(\nu-l)!\binom{\nu}{l}}\,|\nu-l\rangle \\ a^{\dagger k}\,|\nu\rangle &= \sqrt{(\nu+1)\ldots(\nu+k)}\,|\nu+k\rangle = \sqrt{\nu!\binom{\nu+k}{k}}\,|\nu+k\rangle\,. \end{aligned} \tag{X.233}$$

3. Determine $a^{\dagger k}a^l\,|\nu\rangle$ and $a^l a^{\dagger k}\,|\nu\rangle$.
Solution:

$$\begin{aligned} a^{\dagger k}a^l\,|\nu\rangle &= \sqrt{(\nu-l)!\binom{\nu}{l}(\nu-l)!\binom{\nu+k-l}{k}}\,|\nu+k-l\rangle \\ &= \sqrt{\frac{\nu!\,(\nu+k-l)!}{l!k!}}\,|\nu+k-l\rangle \\ a^l a^{\dagger k}\,|\nu\rangle &= \sqrt{\nu!\binom{\nu+k}{k}(\nu+k-l)!\binom{\nu+k}{l}}\,|\nu+k-l\rangle \\ &= \sqrt{\frac{(\nu+k)!\,(\nu+k)!}{l!k!}}\,|\nu+k-l\rangle\,. \end{aligned} \tag{X.234}$$

4. Show that the oscillator length L yields essentially the position of the classical turning points.
Solution: The classical turning points are determined by the equation

$$V = \frac{1}{2} m\omega^2 q_{\text{turning},n}^2 = E_n = \left(n + \frac{1}{2}\right)\hbar\omega. \tag{X.235}$$

This leads to

$$q_{\text{turning},n} = \sqrt{(2n + 1)\frac{\hbar}{m\omega}} = \sqrt{2n + 1} \cdot L. \tag{X.236}$$

For the ground state in particular, we have

$$q_{\text{turning},0} = L. \tag{X.237}$$

5. Proofs by contradiction:

(a) Show by proof of contradiction: There is no largest eigenvalue $\nu_{\max}$.
Solution: If we assume that there is a greatest eigenvalue $\nu_{\max}$, then $a^\dagger |\nu_{\max}\rangle = 0$ must hold. From this it follows directly that:

$$0 = \langle \nu_{\max}| aa^\dagger |\nu_{\max}\rangle = \langle \nu_{\max}| 1 + a^\dagger a |\nu_{\max}\rangle = 1 + \nu_{\max}, \tag{X.238}$$

or $\nu_{\max} = -1$. This is a contradiction since the eigenvalues ν are non-negative. Hence, there is no greatest eigenvalue.

(b) Show by proof of contradiction: The eigenvalues are integers.
Solution: If we assume that the eigenvalues are not integers, then there exists an eigenvector $|\mu\rangle$ with $\hat{n}\,|\mu\rangle = \mu\,|\mu\rangle = (m + \varepsilon)\,|\mu\rangle$; $m \in \mathbb{N}$ and $0 < \varepsilon < 1$. With the notation $a\,|\mu\rangle \equiv |a\mu\rangle$, we obtain

$$\left\langle a^{l+1}\mu \,\middle|\, a^{l+1}\mu\right\rangle = \left\langle a^l\mu \,\middle|\, a^\dagger a \,\middle|\, a^l\mu\right\rangle = \left\langle a^l\mu \,\middle|\, \hat{n} \,\middle|\, a^l\mu\right\rangle. \tag{X.239}$$

With (18.17), it follows that $\hat{n}a^l\,|\mu\rangle = a^l\,(m + \varepsilon - l)\,|\mu\rangle$, and therefore

$$\left\langle a^{l+1}\mu \,\middle|\, a^{l+1}\mu\right\rangle = (m + \varepsilon - l)\left\langle a^l\mu \,\middle|\, a^l\mu\right\rangle. \tag{X.240}$$

If we start from $\left\langle a^l\mu \,\middle|\, a^l\mu\right\rangle = \left\|\left|a^l\mu\right\rangle\right\|^2 \neq 0$, it follows due to $0 < \varepsilon < 1$ that both sides of this equation are not zero for all l. For $l > m + \varepsilon$, the contradiction $\left\|\left|a^{l+1}\mu\right\rangle\right\|^2 < 0$ would follow. Hence we have shown that the assumption $0 < \varepsilon < 1$ is incorrect.

(c) Show that to avoid negative eigenvalues, either (a) the smallest eigenvalue has to be zero or (b) there must be a state $|\nu_{\min}\rangle$ with $a\,|\nu_{\min}\rangle = 0$. Show that in case (b), $\nu_{\min} = 0$.

6. Show that $[q, p] = i\hbar$ (q is the position, p the momentum $\frac{\hbar}{i}\frac{d}{dq}$).
7. Given

$$a := \frac{1}{\sqrt{2\hbar}}\left\{\sqrt{m\omega}q + i\frac{p}{\sqrt{m\omega}}\right\}; \tag{X.241}$$

(a) Derive

$$a^{\dagger} = \frac{1}{\sqrt{2\hbar}} \left\{ \sqrt{m\omega} q - i \frac{p}{\sqrt{m\omega}} \right\}; \tag{X.242}$$

(b) Show that

$$\left[a, a^{\dagger}\right] = 1 \tag{X.243}$$

and

$$H = \hbar\omega \left\{ a^{\dagger} a + \frac{1}{2} \right\}. \tag{X.244}$$

(c) Given the eigenvalue problem

$$\hat{n} \left|\nu\right\rangle = \nu \left|\nu\right\rangle; \; \hat{n} = a^{\dagger} a, \tag{X.245}$$

show that

$$\left\| a \left|\nu\right\rangle \right\|^2 = \nu; \; \left\| a^{\dagger} \left|\nu\right\rangle \right\|^2 = \nu + 1. \tag{X.246}$$

(d) Derive

$$\left[\hat{n}, a\right] = -a; \; \left[\hat{n}, a^{\dagger}\right] = a^{\dagger}. \tag{X.247}$$

(e) Show that

$$\hat{n} a^{l} = a^{l} \left(\hat{n} - l\right); \; \hat{n} a^{\dagger l} = a^{\dagger l} \left(\hat{n} + l\right); \; l = 0, 1, 2, \ldots \tag{X.248}$$

(Proof by mathematical induction.)

(f) Prove

$$\hat{n} a \left|\nu\right\rangle = (\nu - 1)\, a \left|\nu\right\rangle; \; \hat{n} a^{\dagger} \left|\nu\right\rangle = (\nu + 1)\, a^{\dagger} \left|\nu\right\rangle. \tag{X.249}$$

(g) Derive

$$a \left|\nu\right\rangle = \sqrt{\nu} \left|\nu - 1\right\rangle; \; a^{\dagger} \left|\nu\right\rangle = \sqrt{\nu + 1} \left|\nu + 1\right\rangle. \tag{X.250}$$

(h) Show that

$$a^{l} \left|\nu\right\rangle = \sqrt{\nu (\nu - 1) \ldots (\nu - l + 1)} \left|\nu - l\right\rangle. \tag{X.251}$$

X.5 Exercises, Chap. 19

1. Given

$$H = H^{(0)} + F(r)\, \mathbf{l} \cdot \mathbf{s} = \frac{\mathbf{p}^2}{2m} + V(r) + F(r)\, \mathbf{l} \cdot \mathbf{s}. \tag{X.252}$$

(a) Show that:

$$\left[H^{(0)}, l_z\right] = \left[H^{(0)}, s_z\right] = 0; \tag{X.253}$$

(b) Show that:

$$\left[H, l_z\right] \neq 0; \ \left[H, s_z\right] \neq 0; \ \left[H, j_z\right] = 0. \tag{X.254}$$

Hint: See the exercises for Chap. 16.

2. Expand the expression for the relativistic energy levels of the hydrogen atom:

$$E_{nj} = mc^2 \left\{1 + \alpha^2 \left[n - j - \frac{1}{2} + \sqrt{\left(j + \frac{1}{2}\right)^2 - \alpha^2}\right]^{-2}\right\}^{-\frac{1}{2}} - mc^2 \tag{X.255}$$

and compare with the approximation deduced in the text.
Solution: Series expansion with respect to powers of $x \ll 1$ with $x = \alpha^2$ and $y = j + \frac{1}{2}$ yields the expression

$$\begin{aligned} &\left\{1 + x\left[n - y + \sqrt{y^2 - x}\right]^{-2}\right\}^{-\frac{1}{2}} - 1 \\ &= -\frac{1}{2n^2}x + \left(-\frac{1}{2n^3 y} + \frac{3}{8n^4}\right)x^2 + \left(\frac{3}{4n^5 y} - \frac{n + 3y}{8n^4 y3} - \frac{5}{16n^6}\right)x^3 + \cdots \end{aligned} \tag{X.256}$$

3. Given the Hamiltonian

$$H\left|\varphi\right\rangle = \left(H^{(0)} + W\right)\left|\varphi\right\rangle = \left(H^{(0)} + \varepsilon \hat{W}\right)\left|\varphi\right\rangle = E\left|\varphi\right\rangle, \tag{X.257}$$

where the states and the eigenvalues of $H^{(0)}\left|\varphi_n^{(0)}\right\rangle = E_n^{(0)}\left|\varphi_n^{(0)}\right\rangle$ are known (discrete, not degenerate). The initial state is $\left|\varphi_n^{(0)}\right\rangle$ and the corresponding energy is $E_n^{(0)}$. States and energies are expanded in terms of ε

$$\left|\varphi\right\rangle = \left|\varphi_n^{(0)}\right\rangle + \varepsilon\left|\varphi_n^{(1)}\right\rangle + \varepsilon^2\left|\varphi_n^{(2)}\right\rangle + \cdots; \ E = E_n^{(0)} + \varepsilon E_n^{(1)} + \varepsilon^2 E_n^{(2)} + \cdots \tag{X.258}$$

We can assume from the outset that the correction terms are orthogonal to the initial state, $\left\langle\varphi_n^{(0)}\middle|\varphi_n^{(j)}\right\rangle = 0$ for $j \neq 0$. Calculate the corrections to the energy and the state to first order ($\sim\varepsilon^1$, repetition) and to second order ($\sim\varepsilon^2$).

4. We add a perturbation $\sim q^3$ to the Hamiltonian of the harmonic oscillator:

$$H = H^0 + W = -\frac{\hbar^2}{2m}\frac{d^2}{dq^2} + \frac{1}{2}m\omega^2 q^2 + \varepsilon q^3. \tag{X.259}$$

Calculate the correction term of the energy $E_n = \hbar\omega\left(n + \frac{1}{2}\right)$ to first order.

Solution: The correction term for the energy is $\left\langle \varphi_n^{(0)} \right| W \left| \varphi_n^{(0)} \right\rangle$. Since the eigenfunctions of the harmonic oscillator have well-defined parities (see Chap. 18), the integrand is point symmetrical and the integral disappears, $\left\langle \varphi_n^{(0)} \right| W \left| \varphi_n^{(0)} \right\rangle = 0$.

5. Finite nuclear size: For a hydrogen atom, we model the finite core size by the potential

$$V(r) = \begin{cases} -\frac{\gamma}{r} \text{ for } r \geq r_0 \\ \frac{\gamma}{2r_0}\left[\left(\frac{r}{r_0}\right)^2 - 3\right] \text{ for } r \leq r_0 \end{cases} \tag{X.260}$$

(Thus, we replace the point nucleus by a homogenously-charged sphere of radius r_0 with the charge density ρ_0.) Calculate the corrections to the energy in first order. Assume that the radial functions $R_{nl}(r)$ can be approximated for $r \leq r_0$ by $R_{nl}(0)$.

Solution: We have

$$W(r) = \begin{cases} \frac{\gamma}{2r_0}\left[\left(\frac{r}{r_0}\right)^2 - 3 + \frac{2r_0}{r}\right] \text{ for } r \leq r_0 \\ 0 \text{ otherwise.} \end{cases} \tag{X.261}$$

It follows that

$$\left\langle n, l, m \right| W \left| n, l', m' \right\rangle = \int \mathrm{d}\Omega\, Y_l^{m*}(\vartheta, \varphi)\, Y_{l'}^{m'}(\vartheta, \varphi) \times \int_0^\infty r^2 \mathrm{d}r R_{n,l}^*(r) R_{n',l'}(r) W(r). \tag{X.262}$$

This gives

$$\left\langle n, l, m \right| W \left| n, l', m' \right\rangle \approx \delta_{ll'}\delta_{mm'} \cdot \left|R_{n,l}(0)\right|^2 \int_0^{r_0} r^2 \mathrm{d}r W(r) \tag{X.263}$$

or

$$\left\langle n, l, m \right| W \left| n, l', m' \right\rangle \approx \delta_{ll'}\delta_{mm'} \cdot \left|R_{n,l}(0)\right|^2 \int_0^{r_0} r^2 \mathrm{d}r \frac{\gamma}{2r_0}\left[\left(\frac{r}{r_0}\right)^2 - 3 + \frac{2r_0}{r}\right]. \tag{X.264}$$

Evaluation of the integral yields the final result:

$$\left\langle n, l, m \right| W \left| n, l', m' \right\rangle \approx \delta_{ll'}\delta_{mm'} \cdot \left|R_{n,l}(0)\right|^2 \frac{\gamma r_0^2}{10}. \tag{X.265}$$

X.6 Exercises, Chap. 20

1. Given two matrices A and B with

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix};\ B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}. \tag{X.266}$$

   Determine $A \otimes B$.
   Solution:

$$A \otimes B = \begin{pmatrix} 1B & 3B \\ 2B & 1B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 3 & 0 \\ 2 & 1 & 6 & 3 \\ 2 & 0 & 1 & 0 \\ 4 & 2 & 2 & 1 \end{pmatrix}. \tag{X.267}$$

2. Represent the Bell states (20.14) as column vectors. Show in this representation that the Bell states are entangled and that they form a CONS.
   Solution: With

$$|h\rangle \cong \begin{pmatrix} 1 \\ 0 \end{pmatrix};\ |v\rangle \cong \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{X.268}$$

   it follows that

$$|hh\rangle \cong \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix};\ |hv\rangle \cong \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix};\ |vh\rangle \cong \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix};\ |vv\rangle \cong \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{X.269}$$

   and therefore

$$\left|\Psi^{\pm}\right\rangle \cong \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ \pm 1 \\ 0 \end{pmatrix};\ \left|\Phi^{\pm}\right\rangle \cong \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \pm 1 \end{pmatrix}. \tag{X.270}$$

   A column vector of the form $\left(c_1\ c_2\ c_3\ c_4\right)^T$ is factorizable, if it holds that $c_1 \cdot c_4 = c_2 \cdot c_3$; see (20.8). One can see directly that this condition is not satisfied for the Bell states.
   In order to show that the Bell states are a CONS, we consider initially $\left|\Phi^{\pm}\right\rangle$. We have

$$\begin{aligned}&\left|\Phi^+\right\rangle\left\langle\Phi^+\right| + \left|\Phi^-\right\rangle\left\langle\Phi^-\right| \\ &= \frac{1}{2}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}\begin{pmatrix}1 & 0 & 0 & 1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}\begin{pmatrix}1 & 0 & 0 & -1\end{pmatrix}.\end{aligned} \tag{X.271}$$

Multiplication yields

$$\begin{aligned}&\left|\Phi^+\right\rangle\left\langle\Phi^+\right| + \left|\Phi^-\right\rangle\left\langle\Phi^-\right| \\ &= \frac{1}{2}\begin{pmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1&0&0&-1\\0&0&0&0\\0&0&0&0\\-1&0&0&1\end{pmatrix} = \begin{pmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{pmatrix}.\end{aligned} \tag{X.272}$$

Consideration of $\left|\Psi^\pm\right\rangle$ yields the two missing diagonal elements.

3. Two photons are in the state

$$|\Psi\rangle = \frac{|hv\rangle - |vh\rangle}{\sqrt{2}}. \tag{X.273}$$

(a) Show explicitly that it is an entangled state.
(b) Photon 1 passes an analyzer for right-handed circular polarization (the corresponding state reads $\frac{|h\rangle + i|v\rangle}{\sqrt{2}}$). Show that through a measurement, the state $|\Psi\rangle$ is changed into a product state.

4. Show that the Bell states can be transformed into each other by applying the Pauli matrices to a subsystem.
Solution: We transform the system 1, i.e. we apply the operators $\sigma_i \otimes I$. Due to

$$\begin{aligned}\sigma_1 |h\rangle &= |v\rangle \,;\; \sigma_1 |v\rangle = |h\rangle \\ \sigma_2 |h\rangle &= i\,|v\rangle \,;\; \sigma_2 |v\rangle = -i\,|h\rangle \\ \sigma_3 |h\rangle &= |h\rangle \,;\; \sigma_3 |v\rangle = -|v\rangle\,,\end{aligned} \tag{X.274}$$

it follows that e.g.

$$\begin{aligned}(\sigma_1 \otimes I)\left|\Psi^\pm\right\rangle &= \frac{\sigma_1 |h\rangle \otimes |v\rangle \pm \sigma_1 |v\rangle \otimes |h\rangle}{\sqrt{2}} \\ &= \frac{|v\rangle \otimes |v\rangle \pm |h\rangle \otimes |h\rangle}{\sqrt{2}} = \pm\left|\Phi^\pm\right\rangle,\end{aligned} \tag{X.275}$$

and correspondingly for the other Pauli matrices or Bell states (as well as for the operators $I \otimes \sigma_i$).

5. Show that the Bell states are eigenvectors of products of the same Pauli matrices.
Solution: We consider as an example $(\sigma_2 \otimes \sigma_2)\left|\Psi^\pm\right\rangle$. It follows

$$(\sigma_2 \otimes \sigma_2) \left|\Psi^{\pm}\right\rangle = \frac{\sigma_2 \left|h\right\rangle \otimes \sigma_2 \left|v\right\rangle \pm \sigma_2 \left|v\right\rangle \otimes \sigma_2 \left|h\right\rangle}{\sqrt{2}} = \pm \left|\Psi^{\pm}\right\rangle; \tag{X.276}$$

and correspondingly for the other Pauli matrices or Bell states.

6. Transform the inequality (20.27)

$$\cos^2(\alpha - \beta) \leq \cos^2(\alpha - \gamma) + \sin^2(\beta - \gamma) \tag{X.277}$$

for $\alpha = 0$ and $0 < \beta < \pi$ to give

$$\sin(\gamma - \beta)\cos\gamma \leq 0. \tag{X.278}$$

Solution: With $\alpha = 0$, we find

$$\cos^2\beta \leq \cos^2\gamma + \sin^2(\beta - \gamma). \tag{X.279}$$

With the equation $\cos^2 y - \cos^2 x = \sin(x + y)\sin(x - y)$, we then obtain:

$$\cos^2\beta - \cos^2\gamma = \sin(\gamma + \beta)\sin(\gamma - \beta) \leq \sin^2(\gamma - \beta) \tag{X.280}$$

or

$$\sin(\gamma - \beta)\left[\sin(\gamma + \beta) - \sin(\gamma - \beta)\right] \leq 0. \tag{X.281}$$

Using the equation $\sin x - \sin y = 2\cos\frac{x+y}{2}\sin\frac{x-y}{2}$, this gives

$$\sin(\gamma - \beta)\cos\gamma\sin\beta \leq 0. \tag{X.282}$$

For $0 < \beta < \pi$, we have $\sin\beta > 0$, and the inequality

$$\sin(\gamma - \beta)\cos\gamma \leq 0 \tag{X.283}$$

results.

7. Given the function

$$f(\gamma, \beta) = \sin(\gamma - \beta)\cos\gamma; \tag{X.284}$$

determine the position of its zeros and the positions and values of its maxima with respect to γ.
Solution: Zeros exist (a) for $\cos\gamma = 0$, i.e. $\gamma = \frac{\pi}{2} + m_1\pi$ and β arbitrary, as well as for (b) $\sin(\gamma - \beta) = 0$, i.e. $\gamma = \beta + m_2\pi$ with $m_1, m_2 \in \mathbb{Z}$. For the determination of the maxima, we use

$$f(\gamma, \beta) = \frac{\sin(2\gamma - \beta) - \sin\beta}{2}; \quad \frac{\partial f(\gamma, \beta)}{\partial\gamma} = \cos(2\gamma - \beta). \tag{X.285}$$

Hence, we have extrema for $2\gamma - \beta = \frac{\pi}{2} + n\pi$ with $n \in \mathbb{Z}$, whereby there is a maximum/minimum for n even/odd. Inserting this value of γ gives the value for the maxima:

$$f_{\max}(\gamma, \beta) = \frac{1 - \sin\beta}{2}. \tag{X.286}$$

We see that $f_{\max}(\gamma, \beta)$ is always positive for $\beta \neq \frac{\pi}{2}$.

8. A system of two photons is in one of the Bell states. The photon Q1 is incident on an analyzer for horizontal polarization, rotated by an angle α. What is the probability that Q1 passes the analyzer?
Solution: Rotation by the angle α leads to the new horizontally- and vertically-polarized states:

$$|h_\alpha\rangle = \cos\alpha\,|h_0\rangle - \sin\alpha\,|v_0\rangle\,; \quad |v_\alpha\rangle = \sin\alpha\,|h_0\rangle + \cos\alpha\,|v_0\rangle\,. \tag{X.287}$$

First, we obtain Malus' law by evaluating $\langle h_0\,|h_\alpha\rangle$:

$$|\langle h_0\,|h_\alpha\rangle| = \cos^2\alpha. \tag{X.288}$$

The Bell states are

$$\left|\Psi^\pm\right\rangle = \frac{|hv\rangle \pm |vh\rangle}{\sqrt{2}}; \quad \left|\Phi^\pm\right\rangle = \frac{|hh\rangle \pm |vv\rangle}{\sqrt{2}}. \tag{X.289}$$

Measurement of Q1 in the state $|h_{\alpha 1}\rangle$ and comparison with (X.287) yields (abbreviation: $c = \cos\alpha$, $s = \sin\alpha$):

$$\begin{aligned}\langle h_{\alpha 1}\left|\Psi^\pm\right\rangle &= \frac{\langle h_{\alpha 1}\,|h_{01}v_{02}\rangle \pm \langle h_{\alpha 1}\,|v_{01}h_{02}\rangle}{\sqrt{2}} = \frac{c\,|v_{02}\rangle \pm s\,|h_{02}\rangle}{\sqrt{2}} = \frac{|v_{\mp\alpha 2}\rangle}{\sqrt{2}}\\ \langle h_{\alpha 1}\left|\Phi^\pm\right\rangle &= \frac{\langle h_{\alpha 1}\,|h_{01}h_{02}\rangle \pm \langle h_{\alpha 1}\,|v_{01}v_{02}\rangle}{\sqrt{2}} = \frac{c\,|h_{02}\rangle \pm s\,|v_{02}\rangle}{\sqrt{2}} = \frac{|h_{\pm\alpha 2}\rangle}{\sqrt{2}}.\end{aligned} \tag{X.290}$$

We thus have in each case

$$\left|\langle h_{\alpha 1}\left|\Psi^\pm\right\rangle\right|^2 = \left|\langle h_{\alpha 1}\left|\Phi^\pm\right\rangle\right|^2 = \frac{1}{2}. \tag{X.291}$$

Hence, the probability that Q1 passes an arbitrarily adjusted analyzer is 50 % for all cases.

9. Two photons in the state $|h_0\rangle$ are rotated by the angles α and β to give the states $|h_\alpha\rangle$ and $\left|h_\beta\right\rangle$. How does the projection operator referring to $\left|h_\alpha h_\beta\right\rangle$ act on the Bell states?
Solution: The state rotated by the angle α is

$$|h_\alpha\rangle = \cos\alpha\,|h_0\rangle - \sin\alpha\,|v_0\rangle\,. \tag{X.292}$$

We start from the Bell states:

$$\left|\Psi^\pm\right\rangle = \frac{|hv\rangle \pm |vh\rangle}{\sqrt{2}};\ \ \left|\Phi^\pm\right\rangle = \frac{|hh\rangle \pm |vv\rangle}{\sqrt{2}}, \tag{X.293}$$

which we measure in the state

$$\begin{aligned}\left|h_\alpha h_\beta\right\rangle &= [\cos\alpha\,|h_0\rangle - \sin\alpha\,|v_0\rangle]\,[\cos\beta\,|h_0\rangle - \sin\beta\,|v_0\rangle] \\ &= \cos\alpha\cos\beta\,|h_0h_0\rangle - \cos\alpha\sin\beta\,|h_0v_0\rangle - \sin\alpha\cos\beta\,|v_0h_0\rangle \\ &\quad + \sin\alpha\sin\beta\,|v_0v_0\rangle\,.\end{aligned} \tag{X.294}$$

We then have

$$\begin{aligned}\left|h_\alpha h_\beta\right\rangle\left\langle h_\alpha h_\beta\right|\ \Psi^\pm\rangle &= |h_\alpha\beta\rangle\left\langle h_\alpha h_\beta\right|\ \frac{|h_0v_0\rangle \pm |v_0h_0\rangle}{\sqrt{2}} \\ &= -\left|h_\alpha h_\beta\right\rangle \frac{\cos\alpha\sin\beta \pm \sin\alpha\cos\beta}{\sqrt{2}} \\ &= -\left|h_\alpha h_\beta\right\rangle \frac{\sin\left(\beta\pm\alpha\right)}{\sqrt{2}}\end{aligned} \tag{X.295}$$

and

$$\begin{aligned}\left|h_\alpha h_\beta\right\rangle\left\langle h_\alpha h_\beta\right|\ \Phi^\pm\rangle &= \left|h_\alpha h_\beta\right\rangle\left\langle h_\alpha h_\beta\right|\ \frac{|h_0h_0\rangle \pm |v_0v_0\rangle}{\sqrt{2}} \\ &= \left|h_\alpha h_\beta\right\rangle \frac{\cos\alpha\cos\beta \pm \sin\alpha\sin\beta}{\sqrt{2}} \\ &= \left|h_\alpha h_\beta\right\rangle \frac{\cos\left(\beta\mp\alpha\right)}{\sqrt{2}}.\end{aligned} \tag{X.296}$$

10. Given two quantum objects Q1 and Q2, with an N-dimensional CONS $\{|\varphi_i\rangle\}$ for Q1 and $\{|\psi_i\rangle\}$ for Q2 (due to this notation we can omit the index for the number of the quantum object). The initial state is

$$|\chi\rangle = \sum_{ij} c_{ij}\,|\varphi_i\rangle\left|\psi_j\right\rangle. \tag{X.297}$$

What is the probability of measuring Q1 in some state $|\lambda\rangle$ (no matter which state)?
Solution:
Approach 1: The projection of the state $|\chi\rangle$ onto e.g. $|\lambda\rangle\,|\psi_n\rangle$ is

$$\begin{aligned} |\lambda\rangle\,|\psi_n\rangle\,\langle\lambda|\,\langle\psi_n|\,\chi\rangle &= \sum_{ij} c_{ij}\,|\lambda\rangle\,\langle\lambda\,|\varphi_i\rangle\,|\psi_n\rangle\,\langle\psi_n\,\big|\psi_j\big\rangle \\ &= \sum_{i} c_{in}\,\langle\lambda\,|\varphi_i\rangle\,|\lambda\rangle\,|\psi_n\rangle\,. \end{aligned} \tag{X.298}$$

The probability $w(\lambda, \psi_n)$ of measuring the state $|\lambda\rangle\,|\psi_n\rangle$ is thus given by

$$w(\lambda, \psi_n) = \left|\sum_i c_{in}\,\langle\lambda\,|\varphi_i\rangle\right|^2. \tag{X.299}$$

The probability $w(\lambda)$ of measuring Q1 in the state $|\lambda\rangle$ is the sum of the partial probabilities, i.e.

$$w(\lambda) = \sum_n W(\lambda, \psi_n)\,. \tag{X.300}$$

It follows that

$$\begin{aligned} w(\lambda) &= \sum_n \left|\sum_i c_{in}\,\langle\lambda\,|\varphi_i\rangle\right|^2 \\ &= \sum_n \sum_{ij} c_{in}^* c_{jn}\,\langle\varphi_i\,|\lambda\rangle\,\langle\lambda\,\big|\varphi_j\big\rangle \\ &= \langle\lambda|\left(\sum_{ijn} c_{in}^* c_{jn}\,\big|\varphi_j\big\rangle\,\langle\varphi_i|\right)|\lambda\rangle\,. \end{aligned} \tag{X.301}$$

Approach 2: Alternatively, we can deduce $w(\lambda)$ by describing the measurement as

$$|\lambda\rangle\,\langle\lambda|\,\chi\rangle = \sum_{ij} c_{ij}\,|\lambda\rangle\,\langle\lambda\,|\varphi_i\rangle\,\big|\psi_j\big\rangle = \left(\sum_{ij} c_{ij}\,\langle\lambda\,|\varphi_i\rangle\,\big|\psi_j\big\rangle\right)|\lambda\rangle\,. \tag{X.302}$$

The probability of measuring $|\lambda\rangle$ is the squared value of the 'prefactor' of $|\lambda\rangle$, i.e.

$$\begin{aligned} w(\lambda) &= \left|\sum_{ij} c_{ij}\,\langle\lambda\,|\varphi_i\rangle\,\big|\psi_j\big\rangle\right|^2 = \sum_{\text{injm}} c_{in}^*\,\langle\varphi_i\,|\lambda\rangle\,\langle\psi_n|\,c_{jm}\,\langle\lambda\,\big|\varphi_j\big\rangle\,|\psi_m\rangle \\ &= \sum_{\text{injm}} c_{in}^*\,\langle\varphi_i\,|\lambda\rangle\,c_{jm}\,\langle\lambda\,\big|\varphi_j\big\rangle\,\delta_{nm} = \sum_{ijn} c_{in}^* c_{jn}\,\langle\varphi_i\,|\lambda\rangle\,\langle\lambda\,\big|\varphi_j\big\rangle, \end{aligned} \tag{X.303}$$

and we obtain the same result as in (X.301).
Special cases: for $c_{ij} = \delta_{ij} a_i$, we have:

$$|\chi\rangle = \sum_i a_i |\varphi_i\rangle |\psi_i\rangle ; \ w(\lambda) = \langle\lambda| \left(\sum_i |a_i|^2 |\varphi_i\rangle \langle\varphi_i| \right) |\lambda\rangle ; \quad \text{(X.304)}$$

and for $a_i = \frac{e^{i\alpha_i}}{\sqrt{N}}$, we obtain finally

$$|\chi\rangle = \frac{1}{\sqrt{N}} \sum_i |\varphi_i\rangle |\psi_i\rangle ; \ w(\lambda) = \frac{1}{N}. \quad \text{(X.305)}$$

11. Show that entangled states such as the Bell states cannot be 'disentangled' by a reversible transformation of the single-quantum-object basis; entanglement is preserved even in a different basis.
Solution: We begin with the transformation T_i

$$|h_i\rangle = a_i |h'_i\rangle + b_i |v'_i\rangle ; \ |v_i\rangle = c_i |h'_i\rangle + d_i |v'_i\rangle \quad \text{(X.306)}$$

with $a_i d_i \neq b_i c_i$ (in order that T_i^{-1} exists). The four Bell states

$$|\Psi^\pm\rangle = \frac{|hv\rangle \pm |vh\rangle}{\sqrt{2}} ; \ |\Phi^\pm\rangle = \frac{|hh\rangle \pm |vv\rangle}{\sqrt{2}} \quad \text{(X.307)}$$

read in the new basis

$$\begin{aligned} \sqrt{2} |\Psi^\pm\rangle &= [a_1 c_2 \pm c_1 a_2] |h'h'\rangle + [a_1 d_2 \pm c_1 b_2] |h'v'\rangle \\ &+ [b_1 c_2 \pm d_1 a_2] |v'h'\rangle + [b_1 d_2 \pm d_1 b_2] |v'v'\rangle \\ \sqrt{2} |\Phi^\pm\rangle &= [a_1 a_2 \pm c_1 c_2] |h'h'\rangle + [a_1 b_2 \pm c_1 d_2] |h'v'\rangle \\ &+ [b_1 a_2 \pm d_1 c_2] |v'h'\rangle + [b_1 b_2 \pm d_1 d_2] |v'v'\rangle . \end{aligned} \quad \text{(X.308)}$$

The states are factorizable in this basis, if it holds that $a_{h'h'} \cdot a_{v'v'} = a_{h'v'} \cdot a_{v'h'}$. This means that

$$\begin{aligned} [a_1 c_2 \pm c_1 a_2] [b_1 d_2 \pm d_1 b_2] &\stackrel{!}{=} [a_1 d_2 \pm c_1 b_2] [b_1 c_2 \pm d_1 a_2] \\ [a_1 a_2 \pm c_1 c_2] [b_1 b_2 \pm d_1 d_2] &\stackrel{!}{=} [a_1 b_2 \pm c_1 d_2] [b_1 a_2 \pm d_1 c_2] . \end{aligned} \quad \text{(X.309)}$$

Expanding the products and collecting terms leads to

$$\begin{aligned} (a_1 d_1 - c_1 b_1) (c_2 b_2 - d_2 a_2) &\stackrel{!}{=} 0 \\ (a_1 d_1 - c_1 b_1) (a_2 d_2 - b_2 c_2) &\stackrel{!}{=} 0. \end{aligned} \quad \text{(X.310)}$$

This is a contradiction, since we have assumed an invertible ($a_i d_i \neq b_i c_i$) transformation. In other words, the entanglement is 'robust' under such transformations. We know that we can simplify linear combinations of states by a suitable change of basis. In contrast, one can *not* destroy the entanglement.

12. Determine the behavior of the Bell states under reversible transformations. Consider the case of rotations.
Solution: We begin as in exercise 11, with

$$|h_i\rangle = a_i \left|h_i'\right\rangle + b_i \left|v_i'\right\rangle \; ; \; |v_i\rangle = c_i \left|h_i'\right\rangle + d_i \left|v_i'\right\rangle \tag{X.311}$$

and $a_i d_i \neq b_i c_i$ (in order that T_i^{-1} exists). We have the four basis states

$$\left|\Psi^\pm\right\rangle = \frac{|hv\rangle \pm |vh\rangle}{\sqrt{2}} \; ; \; \left|\Phi^\pm\right\rangle = \frac{|hh\rangle \pm |vv\rangle}{\sqrt{2}} \tag{X.312}$$

and want to link them with the transformed states

$$\left|\Psi^\pm\right\rangle' = \frac{\left|h'v'\right\rangle \pm \left|v'h'\right\rangle}{\sqrt{2}} \; ; \; \left|\Phi^\pm\right\rangle' = \frac{\left|h'h'\right\rangle \pm \left|v'v'\right\rangle}{\sqrt{2}}. \tag{X.313}$$

We can again deduce (X.308) and insert in it

$$\begin{array}{l} \left|h'h'\right\rangle = \frac{\left|\Phi^+\right\rangle' + \left|\Phi^-\right\rangle'}{\sqrt{2}} \; ; \; \left|v'v'\right\rangle = \frac{\left|\Phi^+\right\rangle' - \left|\Phi^-\right\rangle'}{\sqrt{2}} \\ \left|h'v'\right\rangle = \frac{\left|\Psi^+\right\rangle' + \left|\Psi^-\right\rangle'}{\sqrt{2}} \; ; \; \left|v'h'\right\rangle = \frac{\left|\Psi^+\right\rangle' - \left|\Psi^-\right\rangle'}{\sqrt{2}}. \end{array} \tag{X.314}$$

It follows that

$$\begin{aligned} 2\left|\Psi^\pm\right\rangle = & [a_1c_2 \pm c_1a_2 + b_1d_2 \pm d_1b_2]\left|\Phi^+\right\rangle' \\ & + [a_1c_2 \pm c_1a_2 - b_1d_2 \mp d_1b_2]\left|\Phi^-\right\rangle' \\ & + [a_1d_2 \pm c_1b_2 + b_1c_2 \pm d_1a_2]\left|\Psi^+\right\rangle' \\ & + [a_1d_2 \pm c_1b_2 - b_1c_2 \mp d_1a_2]\left|\Psi^-\right\rangle' \end{aligned} \tag{X.315}$$

$$\begin{aligned} 2\left|\Phi^\pm\right\rangle = & [a_1a_2 \pm c_1c_2 + b_1b_2 \pm d_1d_2]\left|\Phi^+\right\rangle' \\ & + [a_1a_2 \pm c_1c_2 - b_1b_2 \mp d_1d_2]\left|\Phi^-\right\rangle' \\ & + [a_1b_2 \pm c_1d_2 + b_1a_2 \pm d_1c_2]\left|\Psi^+\right\rangle' \\ & + [a_1b_2 \pm c_1d_2 - b_1a_2 \mp d_1c_2]\left|\Psi^-\right\rangle'. \end{aligned}$$

We specialize to rotations of the single-object basis:

$$a_i = \cos\vartheta_i;\ b_i = -\sin\vartheta_i;\ c_i = -\sin\vartheta_i;\ d_i = \cos\vartheta_i. \tag{X.316}$$

Rearranging the trigonometric functions gives after some manipulations

$$\begin{aligned}
\left|\Psi^+\right\rangle &= \cos\left(\vartheta_1+\vartheta_2\right)\left|\Psi^+\right\rangle' + \sin\left(\vartheta_1+\vartheta_2\right)\left|\Phi^-\right\rangle' \\
\left|\Psi^-\right\rangle &= \cos\left(\vartheta_1-\vartheta_2\right)\left|\Psi^-\right\rangle' - \sin\left(\vartheta_1-\vartheta_2\right)\left|\Phi^+\right\rangle' \\
\left|\Phi^+\right\rangle &= \cos\left(\vartheta_1-\vartheta_2\right)\left|\Phi^+\right\rangle' + \sin\left(\vartheta_1-\vartheta_2\right)\left|\Psi^-\right\rangle' \\
\left|\Phi^-\right\rangle &= \cos\left(\vartheta_1+\vartheta_2\right)\left|\Phi^-\right\rangle' - \sin\left(\vartheta_1+\vartheta_2\right)\left|\Psi^+\right\rangle'.
\end{aligned} \tag{X.317}$$

In particular for $\vartheta_1 = \vartheta_2 = \vartheta$, it follows that

$$\begin{aligned}
\left|\Psi^+\right\rangle &= \cos\left(2\vartheta\right)\left|\Psi^+\right\rangle' + \sin\left(2\vartheta\right)\left|\Phi^-\right\rangle' \\
\left|\Psi^-\right\rangle &= \left|\Psi^-\right\rangle' \\
\left|\Phi^+\right\rangle &= \left|\Phi^+\right\rangle' \\
\left|\Phi^-\right\rangle &= \cos\left(2\vartheta\right)\left|\Phi^-\right\rangle' - \sin\left(2\vartheta\right)\left|\Psi^+\right\rangle'.
\end{aligned} \tag{X.318}$$

X.7 Exercises, Chap. 21

1. Derive the commutation relation (21.26) for position and momentum.
   Solution: We have:

$$U^{-1}(a)\,XU(a) = X + a \text{ with } U = e^{-i\frac{pa}{\hbar}}. \tag{X.319}$$

Expanding in powers of a yields

$$\left(1 + i\frac{Pa}{\hbar}\right)X\left(1 - i\frac{Pa}{\hbar}\right) + O(a^2) = X + a. \tag{X.320}$$

It follows that

$$X + i\frac{Pa}{\hbar}X - Xi\frac{Pa}{\hbar} + O(a^2) = X + a \tag{X.321}$$

or

$$i\frac{P}{\hbar}X - Xi\frac{P}{\hbar} + O(a) = 1 \tag{X.322}$$

and thus for $a \to 0$

$$[X, P] = i\hbar. \tag{X.323}$$

2. Consider the relation between symmetries and conserved quantities by means of the spatial translational invariance of an isolated system of two quantum objects whose interaction depends only on their distance $\mathbf{r}_1 - \mathbf{r}_2$.
   Solution: After shifting by $\mathbf{a}$, the mean value of H for the state $|\varphi\rangle$ must equal the mean value for the shifted state $|\varphi_\mathbf{a}\rangle = e^{-i\frac{\mathbf{pa}}{\hbar}}|\varphi\rangle$:

$$\langle\varphi_\mathbf{a}| H |\varphi_\mathbf{a}\rangle = \langle\varphi| e^{i\frac{\mathbf{pa}}{\hbar}} H e^{-i\frac{\mathbf{pa}}{\hbar}} |\varphi\rangle \stackrel{!}{=} \langle\varphi| H |\varphi\rangle\,. \tag{X.324}$$

   This equation must hold for all $\mathbf{a}$. Then it follows from the last equation that

$$\langle\varphi| 1 + i\frac{\mathbf{pa}}{\hbar}H - Hi\frac{\mathbf{pa}}{\hbar} |\varphi\rangle \stackrel{!}{=} \langle\varphi| H |\varphi\rangle \tag{X.325}$$

   and thus directly

$$\langle\varphi| \mathbf{pa}H - H\mathbf{pa} |\varphi\rangle \stackrel{!}{=} 0. \tag{X.326}$$

   Since this equation has to hold for all $|\varphi\rangle$, it follows that

$$\mathbf{pa}H - H\mathbf{pa} = \mathbf{a}\,(\mathbf{p}H - H\mathbf{p}) = \mathbf{a}\left[\mathbf{p}, H\right] \stackrel{!}{=} 0. \tag{X.327}$$

   Since this equation has to hold for arbitrary $\mathbf{a}$, it follows finally as consequence of spatial translation invariance that

$$\left[H, \mathbf{p}\right] = 0. \tag{X.328}$$

   With this equation, it is also guaranteed that in the series expansion of the e-function in (X.324), all contributions of higher order in $|\mathbf{a}|$ vanish.
3. Let B be a Hermitian operator and U and A a unitary and an antiunitary operator, resp. Show that:

$$e^{iUBU^{-1}} = Ue^{iB}U^{-1};\;\; e^{iABA^{-1}} = Ae^{-iB}A^{-1} \tag{X.329}$$

   Solution: Using the power series representation of the e-function, we have to show that

$$\left(iUBU^{-1}\right)^n = U\,(iB)^n\,U^{-1};\;\; \left(iABA^{-1}\right)^n = A\,(-iB)^n\,A^{-1}. \tag{X.330}$$

   This is done by mathematical induction. Evidently, the statement holds true for $n = 0$.
   In the unitary case, we have

$$\begin{aligned}\left(iUBU^\dagger\right)^{n+1} &= \left(iUBU^\dagger\right)^n iUBU^\dagger = U\,(iB)^n\,U^\dagger iUBU^\dagger \\ &= U\,(iB)^n\,iBU^\dagger = U\,(iB)^{n+1}\,U^\dagger.\end{aligned} \tag{X.331}$$

In the antiunitary case, we find

$$\begin{aligned}\left(iABA^{\dagger}\right)^{n+1} &= \left(iABA^{\dagger}\right)^{n} iABA^{\dagger} = A\left(-iB\right)^{n} A^{\dagger}iABA^{\dagger} = \\ &= A\left(-iB\right)^{n}\left(-iA^{\dagger}\right)ABA^{\dagger} = A\left(-iB\right)^{n}\left(-iB\right)A^{\dagger} \\ &= A\left(-iB\right)^{n+1}A^{\dagger}.\end{aligned} \tag{X.332}$$

4. Show with the help of the propagator U that eigenvalues of A are conserved, if $[H, A] = 0$.
Solution: We proved the statement by means of the Ehrenfest theorem in Chap. 9, Vol. 1. Here we can argue as follows: $[H, A] = 0$ is equivalent to $[U(t - t_0), A] = 0$ (H is independent of the time). It follows with $A|\varphi(t_0)\rangle = a|\varphi(t_0)\rangle$ that

$$A|\varphi(t)\rangle = AU|\varphi(t_0)\rangle = UA|\varphi(t_0)\rangle = Ua|\varphi(t_0)\rangle = a|\varphi(t)\rangle. \tag{X.333}$$

Hence, the eigenvalue a is conserved in time—it is a good quantum number.
5. Consider the translation $\mathbf{r}' = \mathbf{r} + \mathbf{a}$ or $T(\mathbf{a})\,\mathbf{r} = \mathbf{r} + \mathbf{a}$. Show that it can be represented by the unitary transformation $U_{T(\mathbf{a})} = \lim\limits_{n\to\infty}\left(1 - \frac{i}{\hbar}\frac{\mathbf{ap}}{n}\right)^n = e^{-\frac{i}{\hbar}\mathbf{ap}}$.
Solution: For the wavefunctions, we have:

$$\psi(\mathbf{r}) \to \psi'(\mathbf{r});\ \psi'(\mathbf{r}) = U_{T(\mathbf{a})}\psi(\mathbf{r}), \tag{X.334}$$

and it follows that

$$\begin{aligned}\psi'(\mathbf{r} + \mathbf{a}) &= \psi'(T\mathbf{r}) = \psi(\mathbf{r}) \\ \psi'(\mathbf{r}) &= \psi(\mathbf{r} - \mathbf{a}) = U_{T(\mathbf{a})}\psi(\mathbf{r}).\end{aligned} \tag{X.335}$$

U is unitary. In an infinitesimal transformation $\mathbf{r}' = \mathbf{r} + d\mathbf{a}$, it follows that

$$\psi'(\mathbf{r}) = \psi(\mathbf{r} - d\mathbf{a}) = \psi(\mathbf{r}) - d\mathbf{a}\nabla\psi(\mathbf{r}), \tag{X.336}$$

and thus

$$\psi(\mathbf{r}) - d\mathbf{a}\nabla\psi(\mathbf{r}) = U_{T(d\mathbf{a})}\psi(\mathbf{r}). \tag{X.337}$$

We write this in the form

$$U_{T(d\mathbf{a})} = 1 - d\mathbf{a}\nabla = 1 - \frac{i}{\hbar}d\mathbf{ap}. \tag{X.338}$$

In view of the limiting process, we set $d\mathbf{a} = \frac{\mathbf{a}}{n}$ and obtain

$$U_{T(\mathbf{a})} = \lim_{n\to\infty}\left(1 - \frac{i}{\hbar}\frac{\mathbf{ap}}{n}\right)^n = e^{-\frac{i}{\hbar}\mathbf{ap}}. \tag{X.339}$$

6. Determine the commutator of P with an arbitrary function of X, without using $P = \frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}x}$ from the outset (this is to be derived). Use

$$U^{-1}(a)\, X^2 U(a) = U^{-1}(a)\, XU(a)\, U^{-1}(a)\, XU(a) = (X + a)^2\,; \quad \text{(X.340)}$$

and analogously

$$U^{-1}(a)\, X^n U(a) = (X + a)^n \quad \text{(X.341)}$$

as well as the power-series expansion of the function $f(X) = c_0 + c_1 X + c_2 X^2 + \cdots$.

Solution: We have

$$U^{-1}(a)\, XU(a) = X + a \quad \text{(X.342)}$$

as well as

$$\begin{aligned} U^{-1}(a)\, X^n U(a) &= U^{-1}(a)\, X^{n-1} U(a)\, U^{-1}(a)\, XU(a) \\ &= U^{-1}(a)\, X^{n-1} U(a)\,(X + a)\,. \end{aligned} \quad \text{(X.343)}$$

The proposition (X.341) follows by mathematical induction.

Hence, for deducing the commutator we can start from

$$e^{i\frac{Pa}{\hbar}} X^n e^{-i\frac{Pa}{\hbar}} = (X + a)^n\,. \quad \text{(X.344)}$$

For sufficiently small a, it follows that

$$\left(1 + i\frac{Pa}{\hbar}\right) X^n \left(1 - i\frac{Pa}{\hbar}\right) + O(a^2) = X^n + a(n-1)X^{n-1} + O(a^2) \quad \text{(X.345)}$$

and thus

$$\left[P, X^n\right] = \frac{\hbar}{i}(n-1)X^{n-1}. \quad \text{(X.346)}$$

For the commutator with the function $f(X)$, we find:

$$[P, f(X)] = \left[P, \sum c_n X^n\right] = \frac{\hbar}{i}\sum c_n (n-1)X^{n-1} = \frac{\hbar}{i}\frac{\mathrm{d}f(X)}{\mathrm{d}X}. \quad \text{(X.347)}$$

We emphasize that at this point, X and P are still abstract operators.

Analogously, one can derive that

$$[X, f(P)] = i\hbar\frac{\mathrm{d}f(P)}{\mathrm{d}P} \quad \text{(X.348)}$$

holds.

A note in passing: If we choose in particular $f(X) = e^{i\beta X}$, we obtain

$$e^{i\frac{Pa}{\hbar}} e^{i\beta X} e^{-i\frac{Pa}{\hbar}} = e^{i\beta X} e^{i\beta a}$$

i.e. the commutation relation in the *Weyl form*. It is mathematically clearly more well-behaved than the uncertainty principle $[X, P] = i\hbar$, because it contains with $e^{i\beta X}$ etc. only bounded operators (as opposed to X and P, which are not bounded).

7. Show that a rotation through the angle φ around the z axis is represented by $e^{-i\alpha l_z}$.
Solution: We assume a (sufficiently well-behaved) function $f(r, \vartheta, \varphi)$ (spherical coordinates). Taylor expansion yields

$$f(r, \vartheta, \varphi - \alpha) = \sum_n \frac{(-\alpha)^n}{n!} \frac{\partial^n}{\partial \varphi^n} f(r, \vartheta, \varphi). \tag{X.349}$$

With the definition of the z component of the angular momentum:

$$l_z = \frac{\hbar}{i} \frac{\partial}{\partial \varphi}, \tag{X.350}$$

it follows that

$$f(r, \vartheta, \varphi - \alpha) = \sum_n \frac{(-\alpha)^n}{n!} \left(\frac{i}{\hbar}\right)^n l_z^n f(r, \vartheta, \varphi) = e^{-i\alpha l_z} f(r, \vartheta, \varphi). \tag{X.351}$$

The generalization to three-dimensional rotations and abstract angular momentum operators is analogous.

8. Using (21.32)

$$e^{-i\frac{\gamma \mathbf{j}\hat{\mathbf{a}}}{\hbar}} = e^{-i\frac{\varphi j_x}{\hbar}} e^{-i\frac{\gamma j_y}{\hbar}} e^{i\frac{\varphi j_x}{\hbar}}, \tag{X.352}$$

derive the commutation relation for the angular momentum.
Solution: We start with an infinitesimal γ. Then it follows that

$$1 - i\frac{\gamma \mathbf{j}\hat{\mathbf{a}}}{\hbar} = 1 - i e^{-i\frac{\varphi j_x}{\hbar}} \frac{\gamma j_y}{\hbar} e^{i\frac{\varphi j_x}{\hbar}} \text{ or } \mathbf{j}\hat{\mathbf{a}} = e^{-i\frac{\varphi j_x}{\hbar}} j_y e^{i\frac{\varphi j_x}{\hbar}}. \tag{X.353}$$

Because of $\hat{\mathbf{a}} = (0, \cos\varphi, \sin\varphi)$, we can write

$$\mathbf{j}\hat{\mathbf{a}} = \cos\varphi\, j_y + \sin\varphi\, j_z. \tag{X.354}$$

If φ is infinitesimal, we obtain from the last two equations

$$j_y + \varphi\, j_z = \left(1 - i\frac{\varphi j_x}{\hbar}\right) j_y \left(1 + i\frac{\varphi j_x}{\hbar}\right) \tag{X.355}$$

or

$$j_y + \varphi\, j_z = j_y + \varphi \frac{i}{\hbar} j_y j_x - \varphi \frac{i}{\hbar} j_x j_y. \tag{X.356}$$

This leads directly to the commutation relation

$$i\hbar j_z = j_x j_y - j_y j_x = \left[j_x, j_y \right]. \tag{X.357}$$

9. A scalar operator is defined as an operator whose mean value is invariant under a rotation. Derive the equation $[\mathbf{j}, S] = 0$.
Solution: The rotation in $\mathcal{H}$ is $U(\mathcal{R})$, i.e. $|\varphi_{\mathcal{R}}\rangle = U(\mathcal{R})\,|\varphi\rangle$. It must hold that

$$\langle \varphi_{\mathcal{R}}|\, S\, |\varphi_{\mathcal{R}}\rangle = \langle \varphi|\, U^{\dagger}(\mathcal{R}) S U(\mathcal{R})\, |\varphi\rangle = \langle \varphi|\, S\, |\varphi\rangle , \tag{X.358}$$

and thus

$$e^{i\frac{\gamma \mathbf{j}\hat{\mathbf{a}}}{\hbar}} S e^{-i\frac{\gamma \mathbf{j}\hat{\mathbf{a}}}{\hbar}} = S. \tag{X.359}$$

Expansion for infinitesimal γ gives immediately

$$[\mathbf{j}, S] = 0. \tag{X.360}$$

10. A vector operator is an operator $\mathbf{V}$ whose mean value transforms like a vector $\mathbf{v}$ under a rotation through an angle γ about an axis $\hat{\mathbf{a}}$, i.e. as

$$\mathbf{v}' = \cos\gamma \cdot \mathbf{v} + \sin\gamma \cdot \left(\hat{\mathbf{a}} \times \mathbf{v}\right) + (1 - \cos\gamma)\left(\hat{\mathbf{a}} \cdot v\right) \cdot \hat{\mathbf{a}}. \tag{X.361}$$

Derive $[j_i, V_k] = i\hbar \sum_l \varepsilon_{ikl} V_l$.
Solution: The rotation in $\mathcal{H}$ is $U(\mathcal{R})$, i.e. $|\varphi_{\mathcal{R}}\rangle = U(\mathcal{R})\,|\varphi\rangle$. For sufficiently small (infinitesimal) γ, we have

$$\langle \varphi_{\mathcal{R}}|\, V_i\, |\varphi_{\mathcal{R}}\rangle = \langle \varphi|\, U^{\dagger}(\mathcal{R}) V_i U(\mathcal{R})\, |\varphi\rangle = \langle \varphi|\, V_i'\, |\varphi\rangle , \tag{X.362}$$

and accordingly for a rotation:

$$e^{i\frac{\gamma \mathbf{j}\hat{\mathbf{a}}}{\hbar}} V_i e^{-i\frac{\gamma \mathbf{j}\hat{\mathbf{a}}}{\hbar}} = V_i'. \tag{X.363}$$

Let $\hat{\mathbf{a}} = \hat{x}$ and γ be infinitesimal. Then, due to

$$\mathbf{v}' = \mathbf{v} + \gamma \cdot \left(\hat{\mathbf{a}} \times \mathbf{v}\right) + O\left(\gamma^2\right), \tag{X.364}$$

for the transformed vector, it holds that:

$$\mathbf{V}' = \left(V_x, V_y - \gamma V_z, V_z + \gamma V_y\right), \tag{X.365}$$

and this gives for e.g. the y component:

$$\left(1+\frac{i}{\hbar}\gamma j_x\right)V_y\left(1-\frac{i}{\hbar}\gamma j_x\right)=V_y-\gamma V_z, \tag{X.366}$$

i.e. $i\left[j_x, V_y\right]=-\hbar V_z$. Analogously for the other components. In sum, it follows that

$$[j_x, V_x]=0;\ \left[j_x, V_y\right]=i\hbar V_z;\ \left[j_x, V_z\right]=-i\hbar V_y;\ \left[j_i, V_k\right]=i\hbar\sum_l \varepsilon_{ikl}V_l. \tag{X.367}$$

These relations hold e.g. if we insert the position or the momentum for **V**.

11. Formulate explicitly the unitary operator $e^{-i\frac{\gamma}{2}\boldsymbol{\sigma}\hat{\mathbf{a}}}$ for spin 1/2; $\boldsymbol{\sigma}$ is the vector $\boldsymbol{\sigma}=(\sigma_1,\sigma_1,\sigma_1)$ and $\hat{\mathbf{a}}$ a 3-dimensional unit vector.
Solution: We have:

$$e^{-i\frac{\gamma}{2}\boldsymbol{\sigma}\hat{\mathbf{a}}}=\sum_n\frac{1}{n!}\left(-i\frac{\gamma}{2}\boldsymbol{\sigma}\hat{\mathbf{a}}\right)^n. \tag{X.368}$$

With $(\boldsymbol{\sigma}\mathbf{A})(\boldsymbol{\sigma}\mathbf{B})=\mathbf{AB}+i\boldsymbol{\sigma}(\mathbf{A}\times\mathbf{B})$ (already used in the exercises for Chap. 16), it follows that $(\boldsymbol{\sigma}\mathbf{A})^2=\mathbf{A}^2$, or $\left(\boldsymbol{\sigma}\hat{\mathbf{a}}\right)^2=1$, and hence

$$\begin{aligned}
e^{-i\frac{\gamma}{2}\boldsymbol{\sigma}\hat{\mathbf{a}}} &= \sum_{n=0}^{\infty}\frac{1}{(2n)!}(-i)^{2n}\left(\frac{\gamma}{2}\right)^{2n}\left(\boldsymbol{\sigma}\hat{\mathbf{a}}\right)^{2n} \\
&\quad+\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}(-i)^{2n+1}\left(\frac{\gamma}{2}\right)^{2n+1}\left(\boldsymbol{\sigma}\hat{\mathbf{a}}\right)^{2n+1} \\
&= \sum_{n=0}^{\infty}\frac{1}{(2n)!}(-i)^{2n}\left(\frac{\gamma}{2}\right)^{2n}+\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}(-i)^{2n+1}\left(\frac{\gamma}{2}\right)^{2n+1}\left(\boldsymbol{\sigma}\hat{\mathbf{a}}\right) \\
&= \sum_{n=0}^{\infty}\frac{1}{(2n)!}(-1)^n\left(\frac{\gamma}{2}\right)^{2n}-i\left(\boldsymbol{\sigma}\hat{\mathbf{a}}\right)\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}(-1)^n\left(\frac{\gamma}{2}\right)^{2n+1} \\
&= \cos\frac{\gamma}{2}-i\boldsymbol{\sigma}\hat{\mathbf{a}}\sin\frac{\gamma}{2}.
\end{aligned} \tag{X.369}$$

X.8 Exercises, Chap. 22

1. Write the density operator

$$\rho=\sum_n|\varphi_n\rangle\, p_n\,\langle\varphi_n| \tag{X.370}$$

with normalized, but not necessarily orthogonal states $|\varphi_n\rangle$ when it is transformed unitarily.
Solution: The states transform according to

$$\left|\varphi_n'\right\rangle = U\left|\varphi_n\right\rangle \tag{X.371}$$

and we obtain

$$\rho' = U\rho U^\dagger = \sum_n U\left|\varphi_n\right\rangle p_n \left\langle\varphi_n\right| U^\dagger = \sum_n \left|\varphi_n'\right\rangle p_n \left\langle\varphi_n'\right|. \tag{X.372}$$

2. Show that $tr\,(AB) = tr\,(BA)$.
Solution: With the CONS $\{|n\rangle\}$, we find that

$$tr\,(AB) = \sum_n \langle n|AB|n\rangle = \sum_{nm} \langle n|A|m\rangle\langle m|B|n\rangle = \sum_{nm} \langle m|B|n\rangle\langle n|A|m\rangle. \tag{X.373}$$

The last step is possible since $\langle m|B|n\rangle$ and $\langle n|A|m\rangle$ are numbers. Thus we have

$$tr\,(AB) = \sum_{nm} \langle m|B|n\rangle\langle n|A|m\rangle = \sum_{nm} \langle m|BA|m\rangle = tr\,(BA). \tag{X.374}$$

3. Show that the trace is cyclically invariant, i.e.

$$tr\,(ABC) = tr\,(BCA) = tr\,(CAB). \tag{X.375}$$

Solution: Due to $tr\,(AB) = tr\,(BA)$, we can write

$$\begin{aligned} tr\,(ABC) &= tr\,(A\,(BC)) = tr\,((BC)\,A) = tr\,(BCA) \\ tr\,(ABC) &= tr\,((AB)\,C) = tr\,(C\,(AB)) = tr\,(CAB). \end{aligned} \tag{X.376}$$

4. Show that the trace is invariant under unitary transformations.
Solution: With the unitary matrix U, we obtain from the matrix A the new matrix $A' = UAU^{-1}$. Using $tr(AB) = tr(BA)$, it follows that

$$tr\,\left(A'\right) = tr\,\left(UAU^{-1}\right) = tr\,\left(U^{-1}UA\right) = tr\,(A). \tag{X.377}$$

5. Show that the trace is independent of the basis. (This *must* apply, since a basis transformation is unitary.)
Solution: Let two CONS, $\{|\varphi_n\rangle\}$ and $\{|\psi_n\rangle\}$ be given. Then it holds that

$$\begin{aligned} tr\,(A) &= \sum_m \langle\varphi_m|A|\varphi_m\rangle = \sum_{m,n} \langle\varphi_m|\,\psi_n\rangle\langle\psi_n|A|\varphi_m\rangle \\ &= \sum_{m,n} \langle\psi_n|A|\varphi_m\rangle\langle\varphi_m|\,\psi_n\rangle = \sum_n \langle\psi_n|A|\psi_n\rangle. \end{aligned} \tag{X.378}$$

6. Given a CONS $\{|n\rangle\}$ and a state $|\psi\rangle = \sum_n c_n |n\rangle$ with $\sum_n |c_n|^2 = 1$, show that the probability of finding the system in the state m is given by $p_m = tr(\rho |m\rangle \langle m|) = tr(\rho P_m)$.
Solution: The probability sought is given by $|c_m|^2$. Due to $c_m = \langle m| \psi\rangle$, we have:

$$|c_m|^2 = \langle m| \psi\rangle \langle\psi| m\rangle = \langle m| \rho |m\rangle = \sum_n \langle m| n\rangle \langle n| \rho |m\rangle$$
$$= \sum_n \langle n| \rho |m\rangle \langle m| n\rangle = tr\left(\rho |m\rangle \langle m|\right). \tag{X.379}$$

7. Show that for the reduced density operator $\rho^{(1)}$, it holds in general that $tr([\rho^{(1)}]^2) \leq 1$; hence, we have a mixture if the strict inequality applies.
Solution: Being a Hermitian operator, $\rho^{(1)}$ is diagonalizable via a unitary transformation. Hence, there is a diagonal matrix D and a unitary matrix U such that:
$$\rho^{(1)} = UDU^\dagger. \tag{X.380}$$
About the matrix D, we know (since $\rho^{(1)}$ is positive) that its diagonal entries obey $0 \leq d_{nn} \leq 1$ and $tr(D) = 1$. We have (see the previous exercise):
$$tr(\rho^{(1)}) = tr(UDU^\dagger) = tr(D) = 1. \tag{X.381}$$
Moreover, it applies that
$$tr\left(\left[\rho^{(1)}\right]^2\right) = tr\left(UD^2U^\dagger\right) = tr(D^2). \tag{X.382}$$
Due to $0 \leq d_{nn}^2 \leq d_{nn} \leq 1$, it follows that
$$tr\left(\left[\rho^{(1)}\right]^2\right) = tr(D^2) \leq tr(D) = 1. \tag{X.383}$$
The equals sign can apply only if there is just one non-vanishing diagonal element (which, due to $tr\,(\rho) = 1$, must be 1).

8. Write the density operator in the position representation (cf. Chaps. 12 and 13, Vol. 1).
Solution: With
$$\rho = |\psi\rangle \langle\psi|, \tag{X.384}$$
it follows that
$$\langle x| \rho \left|x'\right\rangle = \rho\left(x, x'\right) = \langle x\ |\psi\rangle \langle\psi|\ x'\rangle = \psi(x)\psi^*(x'). \tag{X.385}$$

By applying ρ to a state $|\varphi\rangle$, we obtain an state $|\chi\rangle$; in the abstract and position representations, we have:

$$\begin{aligned} |\chi\rangle &= \rho\,|\varphi\rangle \\ \langle x\,|\chi\rangle &= \chi\,(x) = \int \langle x|\,\rho\,\big|x'\big\rangle\big\langle x'\,|\varphi\rangle\,dx' = \int \rho\left(x, x'\right)\varphi(x')dx'. \end{aligned} \tag{X.386}$$

9. Show explicitly for

$$\rho = \begin{pmatrix} c_1c_1^* & c_1c_2^* \\ c_2c_1^* & c_2c_2^* \end{pmatrix} \tag{X.387}$$

that

$$\rho^2 = \rho \tag{X.388}$$

applies; using this matrix, show explicitly that $\rho^2 = \rho$. Here, it must hold that $|c_1|^2 + |c_2|^2 = 1$.

10. Show that the eigenvalues $\lambda_{1/2}$ of the matrix

$$\rho = \begin{pmatrix} c_1c_1^* & c_1c_2^* \\ c_2c_1^* & c_2c_2^* \end{pmatrix} \tag{X.389}$$

are 0 and 1.
Solution: The eigenvalues of the matrix are calculated by

$$\begin{vmatrix} |c_1|^2 - \lambda & c_1c_2^* \\ c_2c_1^* & |c_2|^2 - \lambda \end{vmatrix} = 0. \tag{X.390}$$

Expansion of the determinant gives

$$\left(|c_1|^2 - \lambda\right)\left(|c_2|^2 - \lambda\right) - c_1c_2^*c_2c_1^* = 0. \tag{X.391}$$

It follows that

$$\lambda^2 - \left(|c_1|^2 + |c_2|^2\right)\lambda + |c_1|^2\,|c_2|^2 - |c_1|^2\,|c_2|^2 = 0. \tag{X.392}$$

Due to $|c_1|^2 + |c_2|^2 = 1$, this yields

$$\lambda^2 - \lambda = 0 \tag{X.393}$$

and thus the proposition is demonstrated.

11. Given the density matrix for a statistical mixture in the form $\rho = p_h\,|h\rangle\,\langle h| + p_v\,|v\rangle\,\langle v|$ or

$$\rho = \begin{pmatrix} p_h & 0 \\ 0 & p_v \end{pmatrix}; \tag{X.394}$$

How does this read in the circularly-polarized basis?
Solution: We have:

$$|h\rangle = \frac{|r\rangle + |l\rangle}{\sqrt{2}}; \quad |v\rangle = \frac{|r\rangle - |l\rangle}{\sqrt{2}i}. \tag{X.395}$$

From this, it follows that:

$$\begin{aligned} \rho = p_h |h\rangle \langle h| + p_v |v\rangle \langle v| &= p_h \frac{|r\rangle + |l\rangle}{\sqrt{2}} \frac{\langle r| + \langle l|}{\sqrt{2}} - p_v \frac{|r\rangle - |l\rangle}{\sqrt{2}i} \frac{\langle r| - \langle l|}{\sqrt{2}i} \\ &= \frac{p_h + p_v}{2} |r\rangle \langle r| + \frac{p_h - p_v}{2} |r\rangle \langle l| + \frac{p_h - p_v}{2} |l\rangle \langle r| + \frac{p_h + p_v}{2} |l\rangle \langle l| , \end{aligned} \tag{X.396}$$

and in matrix form:

$$\rho = \frac{1}{2} \begin{pmatrix} p_h + p_v & p_h - p_v \\ p_h - p_v & p_h + p_v \end{pmatrix}. \tag{X.397}$$

12. Given two quantum objects Q1 and Q2 with the respective N -dimensional CONS $\{|\varphi_i\rangle\}$ for Q1 and $\{|\psi_i\rangle\}$ for Q2 (by the choice of notation, we can omit the index for the number of the quantum object). The initial state is

$$|\chi\rangle = \sum_{ij} c_{ij} |\varphi_i\rangle \left|\psi_j\right\rangle. \tag{X.398}$$

Calculate the probability $w(\lambda)$ of measuring the quantum object 1 in a state $|\lambda\rangle$, and formulate it in terms of the reduced density operator $\rho^{(1)}$.
Solution: In an exercise for Chap. 20, we calculated the probability as

$$w(\lambda) = \sum_{ijn} c_{in}^* c_{jn} \langle \varphi_i |\lambda\rangle \langle \lambda \left|\varphi_j\right\rangle. \tag{X.399}$$

We now consider the density operator $\rho = |\chi\rangle \langle \chi|$; the reduced density operator $\rho^{(1)} = tr_2(\rho)$ is given by:

$$\begin{aligned} \rho^{(1)} &= \sum_n \langle \psi_n |\chi\rangle \langle \chi |\psi_n\rangle \\ &= \sum_n \langle \psi_n| \left(\sum_{ij} c_{ij} |\varphi_i\rangle \left|\psi_j\right\rangle \right) \left(\sum_{kl} c_{kl}^* \langle \varphi_k| \langle \psi_l| \right) |\psi_n\rangle \\ &= \sum_{nijkl} \delta_{nj} c_{ij} |\varphi_i\rangle c_{kl}^* \langle \varphi_k| \delta_{nl} \\ &= \sum_{ijk} c_{ij} |\varphi_i\rangle c_{kj}^* \langle \varphi_k| . \end{aligned} \tag{X.400}$$

By comparison with (X.399), we see that

$$w(\lambda) = \langle\lambda|\, \rho^{(1)}\, |\lambda\rangle\,. \tag{X.401}$$

The probability of measuring the state $|\lambda\rangle$ is thus the expectation value of the reduced density operator, referred to this state.

13. Given the density operator $\rho = \sum_n p_n |\varphi_n\rangle\langle\varphi_n|$, where it holds that $i\hbar\partial_t |\varphi_n\rangle = H|\varphi_n\rangle$. Show that the time behavior of ρ is described by the von-Neumann equation:

$$i\hbar\partial_t\rho = [H, \rho]\,. \tag{X.402}$$

Solution: It holds that

$$\begin{aligned} i\hbar\partial_t\rho &= i\hbar\partial_t \sum_n p_n |\varphi_n\rangle\langle\varphi_n| \\ &= H\sum_n p_n |\varphi_n\rangle\langle\varphi_n| - \sum_n p_n |\varphi_n\rangle\langle\varphi_n| H = [H, \rho]\,. \end{aligned} \tag{X.403}$$

14. Using the example of a polarized photon, show explicitly that a given density matrix does not allow for a unique decomposition.

(a) First formulate the projection operators for the states $|h\rangle$, $|v\rangle$, $|r\rangle$ and $|l\rangle$.
Solution: With $|h\rangle = \begin{pmatrix}1\\0\end{pmatrix}$, $|v\rangle = \begin{pmatrix}0\\1\end{pmatrix}$, $|r\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$ and $|l\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}$, it follows that

$$P_h = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix};\ P_v = \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix} \tag{X.404}$$

and

$$P_r = \frac{1}{2}\begin{pmatrix}1 & -i\\i & 1\end{pmatrix};\ P_l = \frac{1}{2}\begin{pmatrix}1 & i\\-i & 1\end{pmatrix}. \tag{X.405}$$

(b) Given the density matrix $\rho = \frac{1}{2}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$; now formulate the decomposition of ρ in terms of linearly- and circularly-polarized states.
Solution: It evidently applies that $\rho = \frac{1}{2}P_h + \frac{1}{2}P_v$ as well as $\rho = \frac{1}{2}P_r + \frac{1}{2}P_l$. Hence, from a given density matrix, one cannot uniquely determine the underlying states.

15. The spin state of an electron is represented (in the basis of eigenstates of the spin matrix $s_z = \frac{\hbar}{2}\sigma_z$) by the density matrix $\rho = \begin{pmatrix}a & 0\\0 & b\end{pmatrix}$, with $a + b = 1$; $a \geq 0, b \geq 0$.

(a) What is the probability of obtaining $\pm\frac{\hbar}{2}$, if one measures s_x?
Solution: s_x has the eigenvalues $\pm\frac{\hbar}{2}$ and the eigenvectors $|x_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ and $|x_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$. Thereof follow the projectors

$$\begin{aligned}P_1 &= \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1 & 1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\\ P_2 &= \frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix}\begin{pmatrix}1 & -1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix},\end{aligned} \tag{X.406}$$

and thus

$$\begin{aligned}tr\,(\rho P_1) &= tr\left(\begin{pmatrix}a & 0\\0 & b\end{pmatrix}\frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\right)\\ &= \frac{1}{2}tr\begin{pmatrix}a & a\\b & b\end{pmatrix} = \frac{a+b}{2} = \frac{1}{2},\end{aligned} \tag{X.407}$$

as well as

$$\begin{aligned}tr\,(\rho P_2) &= tr\left(\begin{pmatrix}a & 0\\0 & b\end{pmatrix}\frac{1}{2}\begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix}\right)\\ &= \frac{1}{2}tr\begin{pmatrix}a & -a\\-b & b\end{pmatrix} = \frac{a+b}{2} = \frac{1}{2}.\end{aligned} \tag{X.408}$$

Hence, the probability equals $\frac{1}{2}$ for both results.

(b) Calculate the expectation value of s_x and compare it with the trace formalism.
Solution: From the previous part of the exercise, we know that we measure both the eigenvalues $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$ with a probability $\frac{1}{2}$. Accordingly, the mean value is zero. On the other hand, we obtain with $\langle A\rangle = tr\,(\rho A)$:

$$\langle s_x\rangle = tr\left(\begin{pmatrix}a & 0\\0 & b\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\right) = \frac{\hbar}{2}tr\begin{pmatrix}0 & a\\b & 0\end{pmatrix} = 0; \tag{X.409}$$

i.e. agreement.

16. Given a system of two quantum objects; the basis states are in each case $|1\rangle$ and $|2\rangle$.

(a) How is the general total state $|\psi\rangle$ formulated?
Solution:

$$|\psi\rangle = c_{11}\,|1_1 1_2\rangle + c_{12}\,|1_1 2_2\rangle + c_{21}\,|2_1 1_2\rangle + c_{22}\,|2_1 2_2\rangle \quad \text{with} \quad \sum\left|c_{ij}\right|^2 = 1. \tag{X.410}$$

(b) Give explicitly the density matrix for this system.
Solution: We have

$$\rho = \begin{pmatrix} c_{11}c_{11}^* & c_{11}c_{12}^* & c_{11}c_{21}^* & c_{11}c_{22}^* \\ c_{12}c_{11}^* & c_{12}c_{12}^* & c_{12}c_{21}^* & c_{12}c_{22}^* \\ c_{21}c_{11}^* & c_{21}c_{12}^* & c_{21}c_{21}^* & c_{21}c_{22}^* \\ c_{22}c_{11}^* & c_{22}c_{12}^* & c_{22}c_{21}^* & c_{22}c_{22}^*. \end{pmatrix} \tag{X.411}$$

(c) Starting from this matrix, calculate the reduced density matrix $\rho^{(1)}$.
Solution: The unity matrix in space 1 is E_1. It follows that

$$\begin{aligned} A_1 = E_1 \otimes |1_2\rangle \cong \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \\ A_2 = E_1 \otimes |2_2\rangle \cong \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{X.412}$$

The reduced density matrix follows as

$$\rho^{(1)} = A_1^\dagger \rho A_1 + A_2^\dagger \rho A_2. \tag{X.413}$$

Written out, this appears as:

$$\begin{aligned} \rho^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} c_{11}c_{11}^* & c_{11}c_{12}^* & c_{11}c_{21}^* & c_{11}c_{22}^* \\ c_{12}c_{11}^* & c_{12}c_{12}^* & c_{12}c_{21}^* & c_{12}c_{22}^* \\ c_{21}c_{11}^* & c_{21}c_{12}^* & c_{21}c_{21}^* & c_{21}c_{22}^* \\ c_{22}c_{11}^* & c_{22}c_{12}^* & c_{22}c_{21}^* & c_{22}c_{22}^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \\ + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_{11}c_{11}^* & c_{11}c_{12}^* & c_{11}c_{21}^* & c_{11}c_{22}^* \\ c_{12}c_{11}^* & c_{12}c_{12}^* & c_{12}c_{21}^* & c_{12}c_{22}^* \\ c_{21}c_{11}^* & c_{21}c_{12}^* & c_{21}c_{21}^* & c_{21}c_{22}^* \\ c_{22}c_{11}^* & c_{22}c_{12}^* & c_{22}c_{21}^* & c_{22}c_{22}^* \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{X.414}$$

After some calculations, we find

$$\rho^{(1)} = \begin{pmatrix} c_{11}c_{11}^* + c_{12}c_{12}^* & c_{11}c_{21}^* + c_{12}c_{22}^* \\ c_{21}c_{11}^* + c_{22}c_{12}^* & c_{21}c_{21}^* + c_{22}c_{22}^* \end{pmatrix}. \tag{X.415}$$

(d) Show that $tr\left(\rho^{(1)}\right) = 1$ holds.

Solution: From this equation, we read off directly

$$tr\rho^{(1)} = |c_{11}|^2 + |c_{12}|^2 + |c_{21}|^2 + |c_{22}|^2 = 1. \quad \text{(X.416)}$$

The last equals sign holds due to the normalization in (X.410).

(e) Show that $\rho^{(1)} = CC^\dagger$ with $C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$ holds.

Solution:

$$CC^\dagger = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} c_{11}^* & c_{21}^* \\ c_{12}^* & c_{22}^* \end{pmatrix} = \begin{pmatrix} c_{11}c_{11}^* + c_{12}c_{12}^* & c_{11}c_{21}^* + c_{12}c_{22}^* \\ c_{21}c_{11}^* + c_{22}c_{12}^* & c_{21}c_{21}^* + c_{22}c_{22}^* \end{pmatrix}. \quad \text{(X.417)}$$

(f) Calculate $\rho^{(1)2}$.
Solution: We write, abbreviating

$$\rho^{(1)} = \begin{pmatrix} p & a \\ a^* & 1-p \end{pmatrix} \quad \text{(X.418)}$$

and obtain

$$\rho^{(1)2} = \begin{pmatrix} p^2 + |a|^2 & a \\ a^* & (1-p)^2 + |a|^2 \end{pmatrix}. \quad \text{(X.419)}$$

(g) Show that $tr\left(\rho^{(1)2}\right) = 1 - 2\left|\det C\right|^2$ is true.
Solution: We start from

$$\det \rho^{(1)} = p(1-p) - |a|^2 = \det\left(CC^\dagger\right) = \det C \cdot \det C^\dagger = \left|\det C\right|^2. \quad \text{(X.420)}$$

From (X.419), we can read off directly that

$$tr\left(\rho^{(1)2}\right) = 1 - 2p + 2p^2 + 2|a|^2. \quad \text{(X.421)}$$

With $|a|^2 = p(1-p) - \det \rho^{(1)}$, it follows that

$$tr\left(\rho^{(1)2}\right) = 1 - 2\det \rho^{(1)} = 1 - 2\left|\det C\right|^2. \quad \text{(X.422)}$$

17. $\{|\varphi_i\rangle, i = 1, \ldots, N\}$ are normalized, but not necessarily orthogonal states. Show that the density matrix $\rho = \frac{1}{N}\sum_{i=1}^N |\varphi_i\rangle\langle\varphi_i|$ describes a pure state, iff these N states are equal up to a phase.

(a) Let $|\varphi_n\rangle = e^{i\delta_n}|\varphi\rangle$. Show that $\rho^2 = \rho$.
Solution: It follows that $\rho = \frac{1}{N}\sum_{i=1}^N |\varphi\rangle\langle\varphi| = |\varphi\rangle\langle\varphi|$ and $\rho^2 = \rho$.
(b) Let $\rho^2 = \rho$; show that the N states $|\varphi_i\rangle$ are equal up to a phase.
Solution: We introduce a CONS $\{|m\rangle, m = 1, \ldots, N\}$. Then it follows that:

$$\rho = \frac{1}{N}\sum_n |\varphi_n\rangle\,\langle\varphi_n| = \sum_m |m\rangle\, p_m\,\langle m|\,, \tag{X.423}$$

and the probabilities are given by

$$p_m = \langle m|\,\rho\,|m\rangle = \frac{1}{N}\sum_n \langle m\,|\varphi_n\rangle\,\langle\varphi_n|\,m\rangle = \frac{1}{N}\sum_n |\langle m\,|\varphi_n\rangle|^2\,. \tag{X.424}$$

Due to $\rho^2 = \rho$, all p_m have to vanish except one of them, i.e. $p_m = \delta_{mM}$. This means that

$$p_m = \frac{1}{N}\sum_n |\langle m\,|\varphi_n\rangle|^2 = \delta_{mM}\,. \tag{X.425}$$

All the terms in the sum are greater than or equal to zero. Thus, it follows that

$$\begin{matrix} m \neq M \;:\rightarrow\; |\langle m\,|\varphi_n\rangle|^2 = 0\;\forall n \\ m = M \;:\rightarrow\; \frac{1}{N}\sum_n |\langle M\,|\varphi_n\rangle|^2 = 1 \end{matrix}\,. \tag{X.426}$$

Because of $|\langle m\,|\varphi_n\rangle|^2 = 0$ for all $m \neq M$, it follows that all the states $|\varphi_n\rangle$ have to be proportional to $|M\rangle$, i.e. $|\varphi_n\rangle = c_n\,|M\rangle$. Since the $|\varphi_n\rangle$ as well as $|M\rangle$ are normalized, it follows that $|c_n|^2 = 1$ or $c_n = e^{i\alpha_n}\,|M\rangle$. Thus, we have shown that all states are equal, apart from a phase. The second equation in (X.426) is also satisfied, of course, since

$$m = M \;:\rightarrow\; \frac{1}{N}\sum_n |\langle M\,|\varphi_n\rangle|^2 = \frac{1}{N}\sum_n \left|e^{i\alpha_n}\right|^2 = 1. \tag{X.427}$$

X.9 Exercises, Chap. 23

1. Two identical quantum objects are in the states $|\alpha_1\rangle$ and $|\alpha_2\rangle$. Show that the total state must be symmetrical or antisymmetrical,

$$|\psi_\pm\rangle = \frac{|1:\alpha_1, 2:\alpha_2\rangle \pm |1:\alpha_2, 2:\alpha_1\rangle}{\sqrt{2}}. \tag{X.428}$$

Solution: There are initially two equivalent possibilities to describe the product state, namely $|1:\alpha_1, 2:\alpha_2\rangle$ and $|1:\alpha_2, 2:\alpha_1\rangle$. Since these descriptions are indistinguishable (and we cannot exclude one of them), the total state must be a linear combination, i.e.

$$|\Psi\rangle = a\,|1:\alpha_1, 2:\alpha_2\rangle \,+\, b\,|1:\alpha_2, 2:\alpha_1\rangle\,. \tag{X.429}$$

As usual, the state is normalized

$$1 = \langle \Psi \,|\Psi\rangle = |a|^2 + |b|^2 \,. \tag{X.430}$$

Since the product states in (X.429) are equivalent, it follows that

$$a = \frac{e^{i\alpha}}{\sqrt{2}};\; b = \frac{e^{i\beta}}{\sqrt{2}} \tag{X.431}$$

with yet undetermined phases α and β.
Apart from (X.429), there is a further equivalent representation of the total state, namely (commutation of the coefficients a and b)

$$|\Phi\rangle = b\,|1:\alpha_1, 2:\alpha_2\rangle + a\,|1:\alpha_2, 2:\alpha_1\rangle \,. \tag{X.432}$$

Since the states $|\Psi\rangle$ and $|\Phi\rangle$ describe the same facts, they may differ at most by a phase. So it must hold that

$$|\Phi\rangle = e^{i\delta}\,|\Psi\rangle \,. \tag{X.433}$$

Insertion of (X.429) and (X.432) and comparison leads to

$$e^{i\alpha} = e^{i\delta}e^{i\beta};\, e^{i\beta} = e^{i\delta}e^{i\alpha}. \tag{X.434}$$

This yields

$$e^{i\delta} = e^{i(\alpha-\beta)};\quad e^{i\delta} = e^{i(\beta-\alpha)}, \tag{X.435}$$

and, due to $e^{i\delta} = e^{-i\delta}$ or $e^{2i\delta} = 1$, we obtain

$$e^{i\delta} = e^{i(\beta-\alpha)} = \pm 1. \tag{X.436}$$

Thus, it follows that

$$\begin{aligned} |\Psi\rangle &= e^{i\alpha}\frac{|1:\alpha_1, 2:\alpha_2\rangle + e^{i(\beta-\alpha)}\,|1:\alpha_2, 2:\alpha_1\rangle}{\sqrt{2}} \\ &= e^{i\alpha}\frac{|1:\alpha_1, 2:\alpha_2\rangle \pm |1:\alpha_2, 2:\alpha_1\rangle}{\sqrt{2}}. \end{aligned} \tag{X.437}$$

Since the global phase does not play any physical role, we can write finally

$$|\psi_\pm\rangle = \frac{|1:\alpha_1, 2:\alpha_2\rangle \pm |1:\alpha_2, 2:\alpha_1\rangle}{\sqrt{2}}. \tag{X.438}$$

Hence, there are only these two possibilities, bosons (+) and fermions (−).

2. Two identical particles are in the states $|a\rangle$ and $|b\rangle$. What is the correct expression for the total state $|\psi\rangle$?
Solution:

$$|\psi\rangle = \frac{|1:a,2:b\rangle \pm |1:b,2:a\rangle}{\sqrt{2}} \text{ or, written compactly, } \frac{|ab\rangle \pm |ba\rangle}{\sqrt{2}}. \tag{X.439}$$

3. Let $|\varphi\rangle = |1:\alpha_1, 2:\alpha_2, 3:\alpha_3\rangle$. Determine $P_{12}P_{23}|\varphi\rangle$ and $P_{23}P_{12}|\varphi\rangle$. Under what conditions do P_{12} and P_{23} commute?
Solution:

$$\begin{aligned} P_{12}P_{23}|\varphi\rangle &= P_{12}|1:\alpha_1,2:\alpha_3,3:\alpha_2\rangle = |1:\alpha_3,2:\alpha_1,3:\alpha_2\rangle \\ P_{23}P_{12}|\varphi\rangle &= P_{23}|1:\alpha_2,2:\alpha_1,3:\alpha_3\rangle = |1:\alpha_2,2:\alpha_3,3:\alpha_1\rangle . \end{aligned} \tag{X.440}$$

If these two states are to be equal (i.e. they commute), in general it must hold that $\alpha_1 = \alpha_2 = \alpha_3$.

4. Write down explicitly the normalized states $|1:\alpha_1, 2:\alpha_2, \ldots, N:\alpha_N\rangle^{(\pm)}_{norm}$ for 2 and 3 particles.
5. Given 3 identical particles; to save paperwork, we denote the product states simply by $|1,2,3\rangle$ instead of $|1:\alpha_1,2:\alpha_2,3:\alpha_3\rangle$; $|1:\alpha_2,2:\alpha_1,3:\alpha_3\rangle$ is then $|2,1,3\rangle$ etc.

   (a) Write down all 6 product states.
   (b) Show explicitly that for the total (anti)symmetrical state, $P_{12}|\psi\rangle^{\pm} = \eta_{12}|\psi\rangle^{\pm}$. Determine η_{12}.
   (c) Given the state $|\varphi\rangle = |1,2,3\rangle - |1,3,2\rangle + |2,1,3\rangle - |2,3,1\rangle + |3,1,2\rangle - |3,2,1\rangle$, show explicitly that $P_{12}|\varphi\rangle$ can*not* be written as $c|\varphi\rangle$.

6. Show explicitly that $P_{ni}P_{mj}P_{nm}P_{ni}P_{mj} = P_{ij}$.
Solution: We perform each step explicitly by rearranging the state $|\ldots,i:\alpha_i,\ldots,j:\alpha_j,\ldots,m:\alpha_m,\ldots,n:\alpha_n,\ldots\rangle$ correspondingly. This is a bit clumsy and lengthy, but it may help in understanding the problem:

$$\begin{aligned} &P_{mj}\left|\ldots,i:\alpha_i,\ldots,j:\alpha_j,\ldots,m:\alpha_m,\ldots,n:\alpha_n,\ldots\right\rangle \\ &\quad= \left|\ldots,i:\alpha_i,\ldots,j:\alpha_m,\ldots,m:\alpha_j,\ldots,n:\alpha_n,\ldots\right\rangle \\ &P_{ni}\left|\ldots,i:\alpha_i,\ldots,j:\alpha_m,\ldots,m:\alpha_j,\ldots,n:\alpha_n,\ldots\right\rangle \\ &\quad= \left|\ldots,i:\alpha_n,\ldots,j:\alpha_m,\ldots,m:\alpha_j,\ldots,n:\alpha_i,\ldots\right\rangle \\ &P_{nm}\left|\ldots,i:\alpha_n,\ldots,j:\alpha_m,\ldots,m:\alpha_j,\ldots,n:\alpha_i,\ldots\right\rangle \\ &\quad= \left|\ldots,i:\alpha_n,\ldots,j:\alpha_m,\ldots,m:\alpha_i,\ldots,n:\alpha_j,\ldots\right\rangle \\ &P_{mj}\left|\ldots,i:\alpha_n,\ldots,j:\alpha_m,\ldots,m:\alpha_i,\ldots,n:\alpha_j,\ldots\right\rangle \\ &\quad= \left|\ldots,i:\alpha_n,\ldots,j:\alpha_i,\ldots,m:\alpha_m,\ldots,n:\alpha_j,\ldots\right\rangle \\ &P_{ni}\left|\ldots,i:\alpha_n,\ldots,j:\alpha_i,\ldots,m:\alpha_m,\ldots,n:\alpha_j,\ldots\right\rangle \\ &\quad= \left|\ldots,i:\alpha_j,\ldots,j:\alpha_i,\ldots,m:\alpha_m,\ldots,n:\alpha_n,\ldots\right\rangle . \end{aligned} \tag{X.441}$$

We can arrange this in a somewhat clearer manner by using the convention that the quantum numbers for the particles i, j, m, n are always written at the positions $1, 2, 3, 4$, i.e.

$|\ldots, i : \alpha_u, \ldots, j : \alpha_v, \ldots, m : \alpha_w, \ldots, n : \alpha_x, \ldots\rangle \equiv |\alpha_u, \alpha_v, \alpha_w, \alpha_x\rangle$ as an example. Then the calculation reads

$$\begin{aligned}
&\qquad\quad \text{Place} \;\; i \;\;\; j \;\; n \;\; m \\
P_{mj} \left|\alpha_i, \alpha_j, \alpha_m, \alpha_n\right\rangle &= \left|\alpha_i, \alpha_m, \alpha_j, \alpha_n\right\rangle \\
P_{ni} \left|\alpha_i, \alpha_m, \alpha_j, \alpha_n\right\rangle &= \left|\alpha_n, \alpha_m, \alpha_j, \alpha_i\right\rangle \\
P_{nm} \left|\alpha_n, \alpha_m, \alpha_j, \alpha_i\right\rangle &= \left|\alpha_n, \alpha_m, \alpha_i, \alpha_j\right\rangle \\
P_{mj} \left|\alpha_n, \alpha_m, \alpha_i, \alpha_j\right\rangle &= \left|\alpha_n, \alpha_i, \alpha_m, \alpha_j\right\rangle \\
P_{ni} \left|\alpha_n, \alpha_i, \alpha_m, \alpha_j\right\rangle &= \left|\alpha_j, \alpha_i, \alpha_m, \alpha_n\right\rangle .
\end{aligned} \tag{X.442}$$

7. Show that

$$\begin{aligned}
E^{(1)}_{100;nlm} &= \frac{e^2}{4\pi\varepsilon_0} \int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \frac{\left|\psi_{100}(\mathbf{r}_1)\,\psi_{nlm}(\mathbf{r}_2) \pm \psi_{nlm}(\mathbf{r}_1)\,\psi_{100}(\mathbf{r}_2)\right|^2}{2\left|\mathbf{r}_1 - \mathbf{r}_2\right|} \\
&= C_{nl} \pm A_{nl}.
\end{aligned} \tag{X.443}$$

Solution: We have

$$\begin{aligned}
&\int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \frac{\left|\psi_{100}(\mathbf{r}_1)\,\psi_{nlm}(\mathbf{r}_2) \pm \psi_{nlm}(\mathbf{r}_1)\,\psi_{100}(\mathbf{r}_2)\right|^2}{2\left|\mathbf{r}_1 - \mathbf{r}_2\right|} \\
&= \int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \frac{\left[\psi_{100}(\mathbf{r}_1)\,\psi_{nlm}(\mathbf{r}_2) \pm \psi_{nlm}(\mathbf{r}_1)\,\psi_{100}(\mathbf{r}_2)\right]\left[\psi^*_{100}(\mathbf{r}_1)\,\psi^*_{nlm}(\mathbf{r}_2) \pm \psi^*_{nlm}(\mathbf{r}_1)\,\psi^*_{100}(\mathbf{r}_2)\right]}{2\left|\mathbf{r}_1 - \mathbf{r}_2\right|} \\
&= \int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \frac{\left|\psi_{100}(\mathbf{r}_1)\,\psi_{nlm}(\mathbf{r}_2)\right|^2 + \left|\psi_{nlm}(\mathbf{r}_1)\,\psi_{100}(\mathbf{r}_2)\right|^2}{2\left|\mathbf{r}_1 - \mathbf{r}_2\right|} \\
&\quad \pm \int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \frac{\psi_{100}(\mathbf{r}_1)\,\psi_{nlm}(\mathbf{r}_2)\,\psi^*_{nlm}(\mathbf{r}_1)\,\psi^*_{100}(\mathbf{r}_2) + \psi_{nlm}(\mathbf{r}_1)\,\psi_{100}(\mathbf{r}_2)\,\psi^*_{100}(\mathbf{r}_1)\,\psi^*_{nlm}(\mathbf{r}_2)}{2\left|\mathbf{r}_1 - \mathbf{r}_2\right|} \\
&= \int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \frac{\left|\psi_{100}(\mathbf{r}_1)\,\psi_{nlm}(\mathbf{r}_2)\right|^2}{\left|\mathbf{r}_1 - \mathbf{r}_2\right|} \pm \int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \frac{\psi_{100}(\mathbf{r}_1)\,\psi_{nlm}(\mathbf{r}_2)\,\psi^*_{nlm}(\mathbf{r}_1)\,\psi^*_{100}(\mathbf{r}_2)}{\left|\mathbf{r}_1 - \mathbf{r}_2\right|}.
\end{aligned} \tag{X.444}$$

8. Prove (23.71), i.e.

$$w_n = \sum_m \left|^{(-)}\left\langle \varphi_n \psi_m \right| \Phi\Psi\rangle^{(-)}\right|^2 = \left|\left\langle \varphi_n \right| \Phi\rangle\right|^2 . \tag{X.445}$$

Solution: We begin with the relation

$$\begin{aligned}
&4 \sum_m \left|^{(-)}\left\langle \varphi_n \psi_m \right| \Phi\Psi\rangle^{(-)}\right|^2 \\
&= \sum_m \left|\left\langle \varphi_n^{(1)} \right| \Phi^{(1)}\right\rangle\left\langle \psi_m^{(2)} \right| \Psi^{(2)}\right\rangle + \left\langle \varphi_n^{(2)} \right| \Phi^{(2)}\right\rangle\left\langle \psi_m^{(1)} \right| \Psi^{(1)}\right\rangle\right|^2 ,
\end{aligned} \tag{X.446}$$

where we have used (23.68). Written out, this is

$$4\sum_m \left|{}^{(-)}\langle\varphi_n\psi_m\ |\Phi\Psi\rangle^{(-)}\right|^2$$
$$=\sum_m \left\{\begin{array}{l}[\langle 1:\Phi|\ 1:\varphi_n\rangle\langle 2:\Psi|\ 2:\psi_m\rangle + \langle 2:\Phi|\ 2:\varphi_n\rangle\langle 1:\Psi|\ 1:\psi_m\rangle] \\ \cdot[\langle 1:\varphi_n|\ 1:\Phi\rangle\langle 2:\psi_m|\ 2:\Psi\rangle + \langle 2:\varphi_n|\ 2:\Phi\rangle\langle 1:\psi_m|\ 1:\Psi\rangle]\end{array}\right\}. \tag{X.447}$$

We expand and obtain

$$4\sum_m \left|{}^{(-)}\langle\varphi_n\psi_m\ |\Phi\Psi\rangle^{(-)}\right|^2$$
$$=\sum_m \left\{\begin{array}{l}\langle 1:\Phi|\ 1:\varphi_n\rangle\langle 2:\Psi|\ 2:\psi_m\rangle\langle 1:\varphi_n|\ 1:\Phi\rangle\langle 2:\psi_m|\ 2:\Psi\rangle \\ +\langle 1:\Phi|\ 1:\varphi_n\rangle\langle 2:\Psi|\ 2:\psi_m\rangle\langle 2:\varphi_n|\ 2:\Phi\rangle\langle 1:\psi_m|\ 1:\Psi\rangle \\ +\langle 2:\Phi|\ 2:\varphi_n\rangle\langle 1:\Psi|\ 1:\psi_m\rangle\langle 1:\varphi_n|\ 1:\Phi\rangle\langle 2:\psi_m|\ 2:\Psi\rangle \\ +\langle 2:\Phi|\ 2:\varphi_n\rangle\langle 1:\Psi|\ 1:\psi_m\rangle\langle 2:\varphi_n|\ 2:\Phi\rangle\langle 1:\psi_m|\ 1:\Psi\rangle\end{array}\right\}. \tag{X.448}$$

We extract from the sum all the terms which are independent of the summation index:

$$4\sum_m \left|{}^{(-)}\langle\varphi_n\psi_m\ |\Phi\Psi\rangle^{(-)}\right|^2$$
$$=\left\{\begin{array}{c}|\langle 1:\varphi_n|\ 1:\Phi\rangle|^2\sum_m\langle 2:\Psi|\ 2:\psi_m\rangle\langle 2:\psi_m|\ 2:\Psi\rangle \\ +\langle 1:\Phi|\ 1:\varphi_n\rangle\langle 2:\varphi_n|\ 2:\Phi\rangle\sum_m\langle 2:\Psi|\ 2:\psi_m\rangle\langle 1:\psi_m|\ 1:\Psi\rangle \\ +\langle 1:\varphi_n|\ 1:\Phi\rangle\langle 2:\Phi|\ 2:\varphi_n\rangle\sum_m\langle 1:\Psi|\ 1:\psi_m\rangle\langle 2:\psi_m|\ 2:\Psi\rangle \\ +|\langle 2:\varphi_n|\ 2:\Phi\rangle|^2\sum_m\langle 1:\Psi|\ 1:\psi_m\rangle\langle 1:\psi_m|\ 1:\Psi\rangle\end{array}\right\} \tag{X.449}$$

Since $\{|\varphi_n\rangle\}$ and $\{|\psi_m\rangle\}$ are CONS, the terms

$$\begin{array}{l}\sum_m\langle 2:\Psi|\ 2:\psi_m\rangle\langle 2:\psi_m|\ 2:\Psi\rangle = \langle 2:\Psi|\ 2:\Psi\rangle = 1 \\ \sum_m\langle 1:\Psi|\ 1:\psi_m\rangle\langle 1:\psi_m|\ 1:\Psi\rangle = \langle 1:\Psi|\ 1:\Psi\rangle = 1\end{array} \tag{X.450}$$

can be calculated immediately. Due to

$$\langle 1:\psi_m|\ 1:\Psi\rangle = \langle 2:\psi_m|\ 2:\Psi\rangle\,, \tag{X.451}$$

completeness holds also for the two remaining terms, and from this follows the desired result.

X.10 Exercises, Chap. 24

1. Given the density matrix

$$\rho = \begin{pmatrix} |c_1|^2 & c_1 c_2^* e^{i\omega t} \\ c_1^* c_2 e^{-i\omega t} & |c_2|^2 \end{pmatrix}. \tag{X.452}$$

Calculate $\frac{1}{T} \int\limits_0^T \rho \mathrm{d}t$.

Solution: The diagonal elements are clear. For the off-diagonal elements, we have e.g.

$$\frac{1}{T} \int\limits_0^T e^{i\omega t} \mathrm{d}t = \frac{e^{i\omega T} - 1}{i\omega T} = e^{i\omega t T/2} \frac{e^{i\omega T/2} - e^{-i\omega T/2}}{i\omega T} = e^{i\omega T/2} \frac{\sin \omega T/2}{\omega T/2} \tag{X.453}$$

and correspondingly for $e^{-i\omega t}$. The off-diagonal terms tend to zero for $T \to \infty$, and we obtain

$$\frac{1}{T} \int\limits_0^T \rho dt \underset{T\to\infty}{=} \begin{pmatrix} |c_1|^2 & 0 \\ 0 & |c_2|^2 \end{pmatrix}. \tag{X.454}$$

2. Consider the reduced density matrix $\rho_{S,\mathrm{red}} = CC^\dagger$ of (24.16), where C is given as an $M \times N$ matrix:

$$C = (c_{mn}) = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1N} \\ c_{21} & c_{22} & \dots & c_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{M1} & c_{M2} & \dots & c_{MN} \end{pmatrix}. \tag{X.455}$$

Hence, the system has M states, the environment N. Estimate the order of magnitude of the elements of $\rho_{S,\mathrm{red}}$.

Solution: With the N-dimensional row vectors

$$\mathbf{c}_1 = \left(c_{11} \; c_{12} \; c_{13} \dots c_{1N} \right), \tag{X.456}$$

we can write

$$\rho_{S,\mathrm{red}} = CC^\dagger = \begin{pmatrix} \mathbf{c}_1\mathbf{c}_1^\dagger & \mathbf{c}_1\mathbf{c}_2^\dagger & \dots & \mathbf{c}_1\mathbf{c}_M^\dagger \\ \mathbf{c}_2\mathbf{c}_1^\dagger & \mathbf{c}_2\mathbf{c}_2^\dagger & \dots & \mathbf{c}_2\mathbf{c}_M^\dagger \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{c}_M\mathbf{c}_1^\dagger & \mathbf{c}_M\mathbf{c}_2^\dagger & \dots & \mathbf{c}_M\mathbf{c}_M^\dagger \end{pmatrix}. \tag{X.457}$$

Due to $tr\left(\rho_{S,\mathrm{red}}\right) = \sum_{k=1}^M \mathbf{c}_k\mathbf{c}_k^\dagger = 1$, the average value of the diagonal elements is $1/M$. Since each diagonal element consists of N positive summands, we can assume that $\left|c_{jk}\right| = O\left(\frac{1}{\sqrt{M \cdot N}}\right)$. In contrast to the diagonal terms, the off-diagonal elements do not consist of only positive summands; the real and imaginary parts can assume positive and negative values and will (normal distribution provided) cancel for sufficiently large N on average, or add up to zero; and this with the usual statistical error $\sim 1/\sqrt{N}$. Thus we obtain in summary:

$$\left(CC^{\dagger}\right)_{ij} = O\left(\frac{1}{M}\right)\left(\delta_{ij} + O\left(\frac{1}{\sqrt{N}}\right)\right). \tag{X.458}$$

In order to illuminate the argument from another side, we consider the elements $\mathbf{c}_1\mathbf{c}_k^{\dagger}$. Without loss of generality, we position the coordinate system in such a way that $\mathbf{c}_1 = \left(c_{11}\ 0\ 0\ \ldots\ 0\right)$. Then according to the above, we have $c_{11} = O\left(\frac{1}{\sqrt{M}}\right)$ or $\mathbf{c}_1\mathbf{c}_1^{\dagger} = O\left(\frac{1}{M}\right)$. For $k \neq 1$, it holds that $\mathbf{c}_k = \left(c_{k1}\ c_{k2}\ \ldots\ c_{kN}\right)$, whereby the individual components have the average value $\frac{1}{\sqrt{M \cdot N}}$. From this, an estimate follows for the scalar product $\mathbf{c}_1\mathbf{c}_k^{\dagger}$:

$$\mathbf{c}_1\mathbf{c}_k^{\dagger} = c_{11} \cdot c_{k1} = O\left(\frac{1}{\sqrt{M}}\right) \cdot O\left(\frac{1}{\sqrt{M \cdot N}}\right) = O\left(\frac{1}{M\sqrt{N}}\right) \ \text{ for } k \neq 1 \tag{X.459}$$

i.e. again the result (X.458).

3. Calculate explicitly the eigenvalues of the density matrix

$$\rho = \begin{pmatrix} |c_1|^2 & c_1 c_2^* \\ c_1^* c_2 & |c_2|^2 \end{pmatrix} \tag{X.460}$$

with $|c_1| + |c_2|^2 = 1$.
Solution: The eigenvalues are the solutions of the equation

$$\left[|c_1|^2 - \lambda\right]\left[|c_2|^2 - \lambda\right] - |c_1|^2 |c_2|^2 = 0. \tag{X.461}$$

It follows that

$$\lambda^2 - \lambda\left[|c_1| + |c_2|^2\right] = 0 \ \text{ or } \ \lambda^2 - \lambda = 0 \tag{X.462}$$

with the solutions $\lambda = 0, 1$.

4. We consider two quantum objects with $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_1$. The CONS $\{|0\rangle, |1\rangle\}$ is a basis of $\mathcal{H}_1$.

   (a) Show that the states

$$|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}} \tag{X.463}$$

   are also a CONS in $\mathcal{H}_1$.
   Solution: We have

$$\langle \pm | \pm \rangle = \frac{\langle 0| \pm \langle 1|}{\sqrt{2}} \frac{|0\rangle \pm |1\rangle}{\sqrt{2}} = 1; \;\; \langle \pm | \mp \rangle = \frac{\langle 0| \pm \langle 1|}{\sqrt{2}} \frac{|0\rangle \mp |1\rangle}{\sqrt{2}} = 0 \tag{X.464}$$

and

$$\begin{aligned} |+\rangle\langle +| + |-\rangle\langle -| &= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \frac{\langle 0| + \langle 1|}{\sqrt{2}} + \frac{|0\rangle - |1\rangle}{\sqrt{2}} \frac{\langle 0| - \langle 1|}{\sqrt{2}} \\ &= \frac{2|0\rangle\langle 0| + 2|1\rangle\langle 1|}{2} = 1. \end{aligned} \tag{X.465}$$

(b) Write down the states

$$|\psi^{\pm}\rangle = \frac{|01\rangle \pm |10\rangle}{\sqrt{2}} \tag{X.466}$$

in the basis $\{|+\rangle, |-\rangle\}$.
Solution: With

$$|0\rangle = \frac{|+\rangle + |-\rangle}{\sqrt{2}}; \;\; |1\rangle = \frac{|+\rangle - |-\rangle}{\sqrt{2}}, \tag{X.467}$$

we obtain

$$\begin{aligned} |\psi^{\pm}\rangle &= \frac{|01\rangle \pm |10\rangle}{\sqrt{2}} \\ &= \frac{|++\rangle - |+-\rangle + |-+\rangle - |--\rangle}{2\sqrt{2}} \pm \frac{|++\rangle + |+-\rangle - |-+\rangle - |--\rangle}{2\sqrt{2}}, \end{aligned} \tag{X.468}$$

and thus

$$|\psi^{+}\rangle = \frac{|++\rangle - |--\rangle}{\sqrt{2}}; \;\; |\psi^{-}\rangle = \frac{|-+\rangle - |+-\rangle}{\sqrt{2}}. \tag{X.469}$$

(c) As assumed in the text, the effect of the environment is to add to each basis state a corresponding random phase. How are the new states $|\psi^{\pm}\rangle$ formulated?
Solution: For $\{|0\rangle, |1\rangle\}$, it holds that $|0\rangle \rightarrow e^{i\varphi_0}|0\rangle$ and $|1\rangle \rightarrow e^{i\varphi_1}|1\rangle$. It follows that

$$|\psi^{\pm}\rangle \rightarrow e^{i(\varphi_0 + \varphi_1)} \frac{|01\rangle \pm |10\rangle}{\sqrt{2}}. \tag{X.470}$$

For $\{|+\rangle, |-\rangle\}$, we have $|+\rangle \rightarrow e^{i\varphi_+}|+\rangle$ and $|-\rangle \rightarrow e^{i\varphi_-}|-\rangle$. It follows that

$$\begin{aligned} |\psi^{+}\rangle &\rightarrow e^{2i\varphi_+} \tfrac{|++\rangle - e^{2i(\varphi_- - \varphi_+)}|--\rangle}{\sqrt{2}} \\ |\psi^{-}\rangle &\rightarrow e^{i(\varphi_- + \varphi_+)} \tfrac{|-+\rangle - |+-\rangle}{\sqrt{2}}. \end{aligned} \tag{X.471}$$

Only the state $\left|\psi^{+}\right\rangle$ in the basis $\{|+\rangle, |-\rangle\}$ is not decoherence-free under these assumptions.

5. Show that

$$\sum_{i=(m,n)} A_i^{\dagger}(t) A_i(t) = 1;\ A_{i=(m,n)}(t) = \sqrt{p_n}\langle m|\hat{U}(t)|n\rangle. \tag{X.472}$$

See (24.34).
Solution: We have

$$\sum_{i=(m,n)} A_i^{\dagger}(t) A_i(t) = \sum_{m,n} A_{mn}^{\dagger}(t) A_{mn}(t)$$
$$= \sum_{m,n} p_n \langle n|\hat{U}^{\dagger}(t)|m\rangle\langle m|\hat{U}(t)|n\rangle. \tag{X.473}$$

Since the environment states are a CONS, we have $\sum_m |m\rangle\langle m| = 1$. This yields

$$\sum_{i=(m,n)} A_i^{\dagger}(t) A_i(t) = \sum_n p_n \langle n|\hat{U}^{\dagger}(t)\hat{U}(t)|n\rangle = \sum_n p_n = 1. \tag{X.474}$$

6. Two quantum objects each have a two-dimensional Hilbert space with the orthonormal basis vectors $|0\rangle$ and $|1\rangle$. They are in the ground state:

$$|\psi\rangle = c_0|0\rangle|0\rangle + c_1|1\rangle|1\rangle. \tag{X.475}$$

We now perform a change of basis via

$$|0\rangle = a_{11}|+\rangle + a_{12}|-\rangle;\ |1\rangle = a_{21}|+\rangle + a_{22}|-\rangle, \tag{X.476}$$

where $|+\rangle$ and $|-\rangle$ are also an orthonormal basis. Under which conditions does $|\psi\rangle = d_+|+\rangle|+\rangle + d_-|-\rangle|-\rangle$ hold?
Solution: We first note that a change of basis is a unitary transformation:

$$U = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix};\ U^{\dagger} = U^{-1} \tag{X.477}$$

which among other things means

$$\det U = e^{i\delta};\ \delta \in \mathbb{R} \text{ and } a_{22} = \det U \cdot a_{11}^{*};\ a_{21} = -\det U \cdot a_{12}^{*}, \tag{X.478}$$

where we assume in the following $a_{11} \neq 0$ and $a_{12} \neq 0$ in order to exclude trivial cases.
We now insert the basis transformations (X.476) into (X.475) and obtain initially

$$|\psi\rangle = c_0\left(a_{11}^2\,|+\rangle\,|+\rangle \,+\, a_{11}a_{12}\,|+\rangle\,|-\rangle \,+\, a_{12}a_{11}\,|-\rangle\,|+\rangle \,+\, a_{12}^2\,|-\rangle\,|-\rangle\right)$$
$$+\, c_1\left(a_{21}^2\,|+\rangle\,|+\rangle \,+\, a_{21}a_{22}\,|+\rangle\,|-\rangle \,+\, a_{22}a_{21}\,|-\rangle\,|+\rangle \,+\, a_{22}^2\,|-\rangle\,|-\rangle\right). \tag{X.479}$$

According to our premises, the coefficients of $|+\rangle\,|-\rangle$ and $|-\rangle\,|+\rangle$ vanish; thus it follows that:

$$c_0a_{11}a_{12} \,+\, c_1a_{21}a_{22} \stackrel{!}{=} 0. \tag{X.480}$$

With (X.478), this leads to

$$c_0a_{11}a_{12} - c_1e^{2i\delta}\cdot a_{12}^*a_{11}^* \stackrel{!}{=} 0, \tag{X.481}$$

and this means apparently (due to $a_{11}a_{12} \neq 0$) that:

$$|c_0| \stackrel{!}{=} |c_1|\,. \tag{X.482}$$

Hence, the question is answered. But we still want to determine the coefficients of $|+\rangle\,|+\rangle$ and $|-\rangle\,|-\rangle$. We have

$$|\psi\rangle = \left(c_0a_{11}^2 \,+\, c_1a_{21}^2\right)|+\rangle\,|+\rangle \,+\, \left(c_0a_{12}^2 \,+\, c_1a_{22}^2\right)|-\rangle\,|-\rangle\,. \tag{X.483}$$

Because of $a_{21}a_{22} \neq 0$, it follows from (X.480) that:

$$c_1 = -c_0\frac{a_{11}a_{12}}{a_{21}a_{22}} \tag{X.484}$$

and therefore

$$\begin{aligned}|\psi\rangle &= c_0\left(a_{11}a_{22} - a_{12}a_{21}\right)\left[\frac{a_{11}}{a_{22}}\,|+\rangle\,|+\rangle - \frac{a_{12}}{a_{21}}\,|-\rangle\,|-\rangle\right]\\ &= c_0\left[\frac{a_{11}}{a_{11}^*}\,|+\rangle\,|+\rangle\,\acute{} + \frac{a_{12}}{a_{12}^*}\,|-\rangle\,|-\rangle\right];\end{aligned} \tag{X.485}$$

$$d_+ = c_0\frac{a_{11}}{a_{11}^*};\;\; d_- = c_0\frac{a_{12}}{a_{12}^*},\; |d_+| = |d_-|\,. \tag{X.486}$$

X.11 Exercises, Chap. 25

1. Show that:

$$\left|\mathbf{r} - \mathbf{r}'\right| \underset{r\to\infty}{\rightarrow} r - \hat{\mathbf{r}}\cdot\mathbf{r}'. \tag{X.487}$$

Solution:

$$\begin{aligned} \left|\mathbf{r}-\mathbf{r}'\right| = \sqrt{(\mathbf{r}-\mathbf{r}')^2} = \sqrt{r^2 - 2\mathbf{r}\mathbf{r}' + r'^2} = r\sqrt{1 - 2\frac{\mathbf{r}\mathbf{r}'}{r^2} + \frac{r'^2}{r^2}} \\ \left|\mathbf{r}-\mathbf{r}'\right| \underset{r\to\infty}{\to} r\left(1 - \frac{\mathbf{r}\mathbf{r}'}{r^2} + O\left(\frac{1}{r^2}\right)\right) \underset{r\to\infty}{\to} r - \hat{\mathbf{r}}\cdot r'. \end{aligned} \tag{X.488}$$

2. Prove that:

$$\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \underset{r\to\infty}{\to} \frac{e^{ikr}}{r} e^{-i\mathbf{k}'\cdot\mathbf{r}'}. \tag{X.489}$$

Solution: From Exercise 1, it follows that

$$\begin{aligned} \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} &\underset{r\to\infty}{\to} \frac{e^{ik\left(1-\frac{\mathbf{r}\mathbf{r}'}{r^2}+O\left(\frac{1}{r^2}\right)\right)}}{r\left(1-\frac{\mathbf{r}\mathbf{r}'}{r^2}+O\left(\frac{1}{r^2}\right)\right)} \\ &= \frac{e^{ikr}e^{-ik\frac{\mathbf{r}\mathbf{r}'}{r}}e^{ikO\left(\frac{1}{r}\right)}}{r}\left(1+O\left(\frac{1}{r}\right)\right) \underset{r\to\infty}{\to} \frac{e^{ikr}}{r}e^{-ik\hat{\mathbf{r}}r'} = \frac{e^{ikr}}{r}e^{-i\mathbf{k}'\mathbf{r}'}. \end{aligned} \tag{X.490}$$

3. Calculate explicitly the asymptotic form of the current density for the scattered wave.

Solution: With

$$\varphi_{\text{scatt}}(\mathbf{r}) \underset{r\to\infty}{\to} f(\vartheta,\varphi)\frac{e^{ikr}}{r} \tag{X.491}$$

and

$$\mathbf{j} = \frac{\hbar}{2mi}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right), \tag{X.492}$$

it follows asymptotically that:

$$\mathbf{j}_{\text{scatt}} \underset{r\to\infty}{\to} \frac{\hbar}{2mi}\left(f^*(\vartheta,\varphi)\frac{e^{-ikr}}{r}\nabla f(\vartheta,\varphi)\frac{e^{ikr}}{r} - f(\vartheta,\varphi)\frac{e^{ikr}}{r}\nabla f^*(\vartheta,\varphi)\frac{e^{-ikr}}{r}\right). \tag{X.493}$$

For the evaluation we use the representation of the gradient in spherical coordinates (see Appendix D, Vol. 1)

$$\nabla = \frac{\partial}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial}{\partial\vartheta}\mathbf{e}_\vartheta + \frac{1}{r\sin\vartheta}\frac{\partial}{\partial\varphi}\mathbf{e}_\varphi. \tag{X.494}$$

First, we have:

$$\begin{aligned}\nabla f(\vartheta,\varphi)\frac{e^{ikr}}{r} &= f(\vartheta,\varphi)\frac{\partial}{\partial r}\frac{e^{ikr}}{r}\mathbf{e}_r + \frac{e^{ikr}}{r^2}\frac{\partial}{\partial\vartheta}f(\vartheta,\varphi)\,\mathbf{e}_\vartheta \\ &= f(\vartheta,\varphi)\left(ik-\frac{1}{r}\right)\frac{e^{ikr}}{r}\mathbf{e}_r + \frac{e^{ikr}}{r^2}\frac{\partial f(\vartheta,\varphi)}{\partial\vartheta}\mathbf{e}_\vartheta + \frac{e^{ikr}}{r^2\sin\vartheta}\frac{\partial f(\vartheta,\varphi)}{\partial\varphi}\mathbf{e}_\varphi.\end{aligned} \tag{X.495}$$

This leads to

$$\begin{aligned}&f^*(\vartheta,\varphi)\frac{e^{-ikr}}{r}\nabla f(\vartheta,\varphi)\frac{e^{ikr}}{r} \\ &\quad= |f(\vartheta,\varphi)|^2\left(ik-\frac{1}{r}\right)\frac{\mathbf{e}_r}{r^2} + O\left(\frac{1}{r^3}\right)\mathbf{e}_\vartheta + O\left(\frac{1}{r^3}\right)\mathbf{e}_\varphi \rightarrow |f(\vartheta,\varphi)|^2\, ik\frac{\mathbf{e}_r}{r^2},\end{aligned} \tag{X.496}$$

where in the last step we have assumed a sufficiently large r. Then we obtain

$$\mathbf{j}_{\text{scatt}} \underset{r\to\infty}{\rightarrow} \frac{\hbar k}{m}|f(\vartheta,\varphi)|^2\frac{\mathbf{e}_r}{r^2}. \tag{X.497}$$

4. Determine the general relation between scattering amplitude and scattering phases.
Solution: We have for the wavefunction the expression:

$$\varphi(\mathbf{r}) = \sum_{l=0}^{\infty}\frac{u_l(r)}{r}P_l(\cos\vartheta). \tag{X.498}$$

We know that on the other hand that the following expression must hold asymptotically:

$$\varphi_{\text{asy}}(\mathbf{r}) = e^{ikz} + f(\vartheta)\frac{e^{ikr}}{r} = \sum_{l=0}^{\infty}\left[(2l+1)\,i^l j_l(kr) + f_l(\vartheta)\frac{e^{ikr}}{r}\right]P_l(\cos\vartheta). \tag{X.499}$$

Due to the linear independence of the Legendre polynomials, we arrive at

$$\frac{u_l(r)}{r} \rightarrow (2l+1)\,i^l j_l(kr) + f_l(\vartheta)\frac{e^{ikr}}{r}. \tag{X.500}$$

We insert the asymptotic expressions:

$$j_l(kr) \underset{r\to\infty}{\sim} \frac{\sin\left(kr-\frac{l\pi}{2}\right)}{kr};\; u_l(r) \underset{r\to\infty}{\sim} c_l\sin\left(kr-\frac{l\pi}{2}+\delta_l\right), \tag{X.501}$$

and obtain

$$\frac{c_l \sin\left(kr - \frac{l\pi}{2} + \delta_l\right)}{r} = (2l+1)\, i^l \frac{\sin\left(kr - \frac{l\pi}{2}\right)}{kr} + f_l(\vartheta) \frac{e^{ikr}}{r}. \quad \text{(X.502)}$$

We now expand the sine in terms of plane waves

$$c_l \left[e^{i\left(kr - \frac{l\pi}{2} + \delta_l\right)} - e^{-i\left(kr - \frac{l\pi}{2} + \delta_l\right)}\right] = \frac{(2l+1)}{k} i^l \left[e^{i\left(kr - \frac{l\pi}{2}\right)} - e^{-i\left(kr - \frac{l\pi}{2}\right)}\right] + 2i f_l(\vartheta)\, e^{ikr}. \quad \text{(X.503)}$$

Due to the linear independence of the in- and outgoing partial waves, the equations

$$\begin{aligned} c_l e^{i\left(kr - \frac{l\pi}{2} + \delta_l\right)} &= \frac{(2l+1)}{k} i^l e^{i\left(kr - \frac{l\pi}{2}\right)} + 2i f_l(\vartheta)\, e^{ikr} \\ c_l e^{-i\left(kr - \frac{l\pi}{2} + \delta_l\right)} &= \frac{(2l+1)}{k} i^l e^{-i\left(kr - \frac{l\pi}{2}\right)} \end{aligned} \quad \text{(X.504)}$$

must hold. The second equation yields

$$c_l = \frac{(2l+1)}{k} i^l e^{i\delta_l}, \quad \text{(X.505)}$$

and it follows that:

$$f_l(\vartheta) = \frac{(2l+1)}{2ik}\left(e^{2i\delta_l} - 1\right) = \frac{(2l+1)}{k} e^{i\delta_l} \sin\delta_l. \quad \text{(X.506)}$$

Hence, the connection between scattering amplitude and scattering phases is

$$f(\vartheta) = \sum_{l=0}^{\infty} f_l(\vartheta)\; P_l(\cos\vartheta) = \frac{1}{k}\sum_{l=0}^{\infty} (2l+1)\, e^{i\delta_l} \sin\delta_l\, P_l(\cos\vartheta). \quad \text{(X.507)}$$

5. Determine the radial equations for a general potential $V(\mathbf{r})$.
Solution: We start from the SEq in the form

$$\left(\nabla^2 + k^2 - v_{\text{eff}}(\mathbf{r})\right) \psi(\mathbf{r}) = 0. \quad \text{(X.508)}$$

The multipole expansions of the potential and the wavefunction read (we use here the abbreviation $\hat{\mathbf{r}} = \vartheta, \varphi$)

$$v_{\text{eff}}(\mathbf{r}) = \sum_{lm} w_{lm}(r)\, Y_l^m\left(\hat{\mathbf{r}}\right); \; \psi(\mathbf{r}) = \sum_{lm} \frac{u_{lm}(r)}{r} Y_l^m\left(\hat{\mathbf{r}}\right). \quad \text{(X.509)}$$

Insertion in the SEq yields

$$\sum_{lm}\left(\frac{d^2}{dr^2}-\frac{l(l+1)}{r^2}+k^2\right)u_{lm}(r)\,Y_l^m(\hat{\mathbf{r}})-\sum_{l_1m_1}w_{l_1m_1}(r)\,Y_{l_1}^{m_1}(\hat{\mathbf{r}})$$
$$\times\sum_{l_2m_2}u_{l_2m_2}(r)\,Y_{l_2}^{m_2}(\hat{\mathbf{r}})=0. \tag{X.510}$$

We transform the last expression by means of (cf. Chap. 16 and Appendix B, Vol. 2):

$$Y_{l_1}^{m_1}(\hat{\mathbf{r}})\,Y_{l_2}^{m_2}(\hat{\mathbf{r}})=\sum_{L=|l_1-l_2|}^{l_1+l_2}\sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2L+1)}}\langle l_1l_200|L0\rangle\langle l_1l_2m_1m_2|L\,m_1+m_2\rangle Y_L^M(\hat{\mathbf{r}}), \tag{X.511}$$

and obtain initially

$$0=\sum_{lm}\left(\frac{d^2}{dr^2}-\frac{l(l+1)}{r^2}+k^2\right)u_{lm}(r)\,Y_l^m(\hat{\mathbf{r}})$$
$$-\sum_{l_1m_1l_2m_2L}\sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2L+1)}}\langle l_1l_200|L0\rangle\langle l_1l_2m_1m_2|L\,m_1+m_2\rangle w_{l_1m_1}(r)u_{l_2m_2}(r)\,Y_L^M(\hat{\mathbf{r}})\,. \tag{X.512}$$

Due to the orthogonality of the spherical harmonics ($\int\left[Y_l^m(\hat{r})\right]^* Y_{l'}^{m'}(\hat{r})\,d\Omega=\delta_{ll'}\delta_{mm'}$), we obtain from this:

$$0=\left(\frac{d^2}{dr^2}-\frac{l(l+1)}{r^2}+k^2\right)u_{lm}(r)$$
$$-\sum_{l_1m_1}\sum_{l_2=|l-l_1|}^{l+l_1}\sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}}\langle l_1l_200|l0\rangle\langle l_1l_2m_1\,m-m_1|l\,m\rangle w_{l_1m_1}(r)\,u_{l_2m-m_1}(r)\,. \tag{X.513}$$

We see that in the second term, there are radial functions coupled for different angular momentum indices.

6. The Yukawa potential (also called the screened Coulomb potential) has the form

$$V(r)=V_0\frac{e^{-r/a}}{r};\ a>0. \tag{X.514}$$

The range of the potential is of order a. Determine the scattering amplitude for the potential in the Born approximation. The Coulomb potential follows for $a\to\infty$ (infinite range of the Coulomb potential). Calculate also in this case the scattering cross section (*Rutherford scattering cross section*).
Solution: The scattering amplitude in the Born approximation is given by

$$f_{\text{Born}}(\vartheta, \varphi) = -\frac{2m}{q\hbar^2} \int_0^\infty \mathrm{d}r\; rV(r) \sin qr$$
$$= -\frac{2mV_0}{q\hbar^2} \int_0^\infty \mathrm{d}r\; e^{-r/a} \sin qr = -\frac{2mV_0}{\hbar^2} \frac{a^2}{1 + a^2q^2}. \tag{X.515}$$

With $q = 2k \sin \frac{\vartheta}{2}$, it follows that:

$$f_{\text{Born}}(\vartheta, \varphi) = -\frac{2mV_0}{\hbar^2} \frac{a^2}{1 + 4a^2k^2 \sin^2 \frac{\vartheta}{2}} \tag{X.516}$$

or

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}_{\text{Born}} = \left(\frac{2mV_0}{\hbar^2}\right)^2 \left(\frac{a^2}{1 + 4a^2k^2 \sin^2 \frac{\vartheta}{2}}\right)^2. \tag{X.517}$$

The Coulomb potential follows in the limit $a \to \infty$ (infinite range of the Coulomb potential), and the scattering cross section in this case is given by:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}_{\text{Born}} = \left(\frac{2mV_0}{\hbar^2}\right)^2 \frac{1}{16k^4 \sin^4 \frac{\vartheta}{2}}^2. \tag{X.518}$$

With

$$E = \frac{\hbar^2k^2}{2m}, V_0 = q^2, \tag{X.519}$$

we find the usual form of the Rutherford scattering cross section:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}_{\text{Born}} = \left(\frac{2mq^2}{\hbar^2}\right)^2 \left(\frac{\hbar^2}{2mE}\right)^2 \frac{1}{16 \sin^4 \frac{\vartheta}{2}} = q^4 \frac{1}{16E^2 \sin^4 \frac{\vartheta}{2}}. \tag{X.520}$$

7. In this exercise, we address the transformation between the abstract representation and the position representation. We recall that this topic is discussed in more detail in Chap. 12, Vol. 1.

   (a) Transform the equation

$$|\psi\rangle = |\psi_0\rangle + Gv\,|\psi\rangle \tag{X.521}$$

   into the position representation.
   Solution: We multiply by $\langle \mathbf{r}|$ from the left and insert the 1 (i.e. $\int \mathrm{d}^3r' \left|\mathbf{r}'\right\rangle\left\langle\mathbf{r}'\right|$) twice on the right-hand side. This gives

$$\langle \mathbf{r} | \psi \rangle = \langle \mathbf{r} | \psi_0 \rangle + \int d^3r' \int d^3r'' \langle \mathbf{r} | G | \mathbf{r}' \rangle \langle \mathbf{r}' | v | \mathbf{r}'' \rangle \langle \mathbf{r}'' | \psi \rangle . \quad \text{(X.522)}$$

The potential operator v is local:

$$\langle \mathbf{r}' | v | \mathbf{r}'' \rangle = \langle \mathbf{r}' | v | \mathbf{r}'' \rangle \delta \left(\mathbf{r}' - \mathbf{r}'' \right) = v \left(\mathbf{r}' \right) \delta \left(\mathbf{r}' - \mathbf{r}'' \right) , \quad \text{(X.523)}$$

and with

$$\langle \mathbf{r} | \psi \rangle = \psi (\mathbf{r}) ; \ \langle \mathbf{r} | G | \mathbf{r}' \rangle = -\frac{1}{4\pi} \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \quad \text{(X.524)}$$

we obtain

$$\psi (\mathbf{r}) = \psi_0 (\mathbf{r}) - \frac{1}{4\pi} \int d^3r' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} v \left(\mathbf{r}' \right) \psi \left(\mathbf{r}' \right) . \quad \text{(X.525)}$$

(b) Write the right-hand side of the following equation:

$$f_{\text{Born}} (\vartheta, \varphi) = -\frac{1}{4\pi} \langle \mathbf{k}' | v | \mathbf{k} \rangle \quad \text{(X.526)}$$

explicitly in the position representation.
Solution: On the right-hand side, we insert the 1 (i.e. $\int d^3r \, |\mathbf{r}\rangle \langle \mathbf{r}|$) twice:

$$f_{\text{Born}} (\vartheta, \varphi) = -\frac{1}{4\pi} \int d^3r' \int d^3r'' \langle \mathbf{k}' | \mathbf{r}' \rangle \langle \mathbf{r}' | v | \mathbf{r}'' \rangle \langle \mathbf{r}'' | \mathbf{k} \rangle . \quad \text{(X.527)}$$

We then have

$$\langle \mathbf{r} | \mathbf{k} \rangle = e^{i\mathbf{kr}} ; \ \langle \mathbf{k} | \mathbf{r} \rangle = \langle \mathbf{r} | \mathbf{k} \rangle^{\dagger} = \langle \mathbf{r} | \mathbf{k} \rangle^{*} = e^{-i\mathbf{kr}} \quad \text{(X.528)}$$

(recall that here, we omit the normalization factor $(2\pi)^{-3/2}$, in contrast to Chap. 12, Vol. 1). Since the potential operator v is local, it follows that:

$$f_{\text{Born}} (\vartheta, \varphi) = -\frac{1}{4\pi} \int d^3r' \, e^{-i\mathbf{k}'\mathbf{r}'} v \left(\mathbf{r}' \right) e^{i\mathbf{kr}'} . \quad \text{(X.529)}$$

With $v (\mathbf{r}) = \frac{2m}{\hbar^2} V (\mathbf{r})$ and $\mathbf{q} := \mathbf{k} - \mathbf{k}'$, we arrive at

$$f_{\text{Born}} (\vartheta, \varphi) = -\frac{m}{2\pi\hbar^2} \int d^3r' \, V \left(\mathbf{r}' \right) e^{i\mathbf{qr}'} . \quad \text{(X.530)}$$

X.12 Exercises, Chap. 26

1. Above, it was proposed that you yourself try to find the prime factorization of 268898680104636581 and 170699960169639253. Did you find it?
   If not, here is the solution: The result is $268898680104636581 = 998653 \cdot 998681 \cdot 269617$ and $170699960169639253 = 413158511 \cdot 413158523$.
2. Pauli matrices and qubits

   (a) How do the Pauli matrices act on the qubit states $|0\rangle$ and $|1\rangle$?
   Solution: In the following we do not distinguish between $\cong$ and $=$. We have:

$$\begin{aligned} \sigma_x |0\rangle &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle \\ \sigma_x |1\rangle &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle ; \end{aligned} \tag{X.531}$$

   and analogously:

$$\begin{aligned} &\sigma_y |0\rangle = i\,|1\rangle ; \;\; \sigma_y |1\rangle = -i\,|0\rangle \\ &\sigma_z |0\rangle = |0\rangle ; \;\; \sigma_z |1\rangle = -|1\rangle . \end{aligned} \tag{X.532}$$

   (b) How do the Pauli matrices act on the qubit state $|\varphi\rangle = c\,|0\rangle + d\,|1\rangle$?
   Solution: With the results of Part (a), we obtain:

$$\begin{aligned} \sigma_x |\varphi\rangle &= d\,|0\rangle + c\,|1\rangle \\ \sigma_y |\varphi\rangle &= -id\,|0\rangle + ic\,|1\rangle \\ \sigma_z |\varphi\rangle &= c\,|0\rangle - d\,|1\rangle . \end{aligned} \tag{X.533}$$

3. Calculate the full expression containing N terms:

$$|z\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \cdots \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{X.534}$$

   Solution: We consider first as an example the case $N = 3$. Expanding yields:

$$|z\rangle = \frac{|000\rangle - |001\rangle - |010\rangle + |011\rangle - |100\rangle + |101\rangle + |110\rangle - |111\rangle}{2\sqrt{2}}. \tag{X.535}$$

   We see that in the binary representation, the sign of $|a_1 a_2 a_3\rangle$ is given by $(-1)^{a_1 + a_2 + a_3}$. Generalizing, we can conclude that for an arbitrary natural number k (in any representation, whereby we confine ourselves here to decimal numbers), the sign $(-1)^{t_k}$ can be calculated as follows: We represent k as a binary number. If the number of 1's is odd, then $t_k = 1$; if it is even, we have $t_k = 0$. (For this reason, the numbers with $t_k = 1$ and $t_k = 0$ are sometimes called *od*ious numbers and *ev*il numbers.) For a general state of the form (X.534), we have in the decimal representation:

$$|z\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \cdots \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} = \frac{1}{2^{N/2}} \sum_{k=0}^{2^N-1} (-1)^{t_k} |k\rangle . \quad \text{(X.536)}$$

Remark: The series of the t_k begins with

$$01101001100101101001011001101001 \ldots . \quad \text{(X.537)}$$

It is called the *Thue–Morse series* and occurs in different areas (number theory, combinatorics, fractals, computer-generated music, etc.).[173]

4. Show that:

$$|q\rangle \to H|q\rangle = \frac{|1-q\rangle + (-1)^q |q\rangle}{\sqrt{2}}; \; q \in \{0, 1\} \quad \text{(X.538)}$$

where H is the Hadamard matrix.
Solution: With

$$|0\rangle \cong \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \;\; |1\rangle \cong \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{(X.539)}$$

it follows that:

$$\begin{aligned} H|0\rangle &\cong \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\ H|1\rangle &\cong \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \end{aligned} \quad \text{(X.540)}$$

One can combine the last two equations in various ways; one possibility is

$$H|q\rangle = \frac{|1-q\rangle + (-1)^q |q\rangle}{\sqrt{2}}; \; q \in \{0, 1\} . \quad \text{(X.541)}$$

5. Calculate explicitly

$$\Phi_\varphi H \Phi_\vartheta H \quad \text{(X.542)}$$

where H is the Hadamard transformation and Φ the phase shifter.
Solution: We first use the matrix representation:

$$H \cong \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}; \; \Phi_\varphi \cong \begin{pmatrix} 1 & 0 \\ 0 & e^{i\varphi} \end{pmatrix}. \quad \text{(X.543)}$$

[173] 'Construction principle': Starting with 0, we replace in every step a 0 by 01 and a 1 by 10. Thus we obtain: $0 \to 01 \to 0110 \to 01101001 \to \cdots$.

This gives

$$\Phi_\varphi H \Phi_\vartheta H = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & e^{i\varphi} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & e^{i\vartheta} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$
$$= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ e^{i\varphi} & -e^{i\varphi} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ e^{i\vartheta} & -e^{i\vartheta} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+e^{i\vartheta} & 1-e^{i\vartheta} \\ e^{i\varphi}\left(1-e^{i\vartheta}\right) & e^{i\varphi}\left(1+e^{i\vartheta}\right) \end{pmatrix}. \tag{X.544}$$

By extracting the factor $e^{i\vartheta/2}e^{i\varphi/2}$, we can write (X.544) as

$$\Phi_\varphi H \Phi_\vartheta H = e^{i\vartheta/2}\begin{pmatrix} \cos\frac{\vartheta}{2} & -i\sin\frac{\vartheta}{2} \\ -ie^{i\varphi}\sin\frac{\vartheta}{2} & e^{i\varphi}\cos\frac{\vartheta}{2}. \end{pmatrix} \tag{X.545}$$

Another possibility is offered by the relation

$$|q\rangle \to H|q\rangle = \frac{|1-q\rangle + (-1)^q|q\rangle}{\sqrt{2}};\ |q\rangle \to \Phi_\varphi|q\rangle = e^{iq\varphi}|q\rangle;\ q \in \{0,1\}. \tag{X.546}$$

From it, we find a representation which is completely equivalent to (X.544):

$$\begin{aligned}\Phi_\varphi H \Phi_\vartheta H\,|q\rangle &= \Phi_\varphi H \Phi_\vartheta \frac{|1-q\rangle + (-1)^q\,|q\rangle}{\sqrt{2}} = \Phi_\varphi H \frac{e^{i(1-q)\vartheta}\,|1-q\rangle + (-1)^q\,e^{iq\vartheta}\,|q\rangle}{\sqrt{2}} \\ &= \Phi_\varphi \frac{\left[e^{i(1-q)\vartheta} + e^{iq\vartheta}\right]|q\rangle + (-1)^{1-q}\left[e^{i(1-q)\vartheta} - e^{iq\vartheta}\right]|1-q\rangle}{2} \\ &= \frac{e^{iq\varphi}\left[e^{i(1-q)\vartheta} + e^{iq\vartheta}\right]|q\rangle + (-1)^{1-q}\,e^{i(1-q)\varphi}\left[e^{i(1-q)\vartheta} - e^{iq\vartheta}\right]|1-q\rangle}{2} \\ &= \frac{e^{iq\varphi}\left[1 + e^{i\vartheta}\right]|q\rangle + e^{i(1-q)\varphi}\left[1 - e^{i\vartheta}\right]|1-q\rangle}{2} \\ &= e^{i\vartheta/2}\left(e^{iq\varphi}\cos\frac{\vartheta}{2}\,|q\rangle - ie^{i(1-q)\varphi}\sin\frac{\vartheta}{2}\,|1-q\rangle\right).\end{aligned} \tag{X.547}$$

6. Kickback and Grover's algorithm: Given that

$$f(k) = \delta_{k\kappa};\ k = 0, 1, \ldots, d-1;\ d = 2^n;\ 0 \leq \kappa \leq d-1. \tag{X.548}$$

The effect of the kickback may be written as:

$$|k\rangle \to (-1)^{f(k)}\,|k\rangle \ \text{ or } U_\kappa\,|k\rangle = (-1)^{f(k)}\,|k\rangle\,, \tag{X.549}$$

where $\{|k\rangle\}$ is a CONS. Show that:

$$U_\kappa = 1 - 2\,|\kappa\rangle\,\langle\kappa|\,. \tag{X.550}$$

Solution: It holds that

$$U_\kappa\,|k\rangle\,\langle k| = (-1)^{f(k)}\,|k\rangle\,\langle k|\,. \tag{X.551}$$

We sum over the states, use $\sum |k\rangle\,\langle k| = 1$, and obtain:

$$\begin{aligned} U_\kappa &= \sum_{k=0}^{d-1} (-1)^{f(k)}\,|k\rangle\,\langle k| = \sum_{k=0,\neq\kappa}^{d-1} |k\rangle\,\langle k| - |\kappa\rangle\,\langle\kappa| \\ &= \sum_{k=0}^{d-1} |k\rangle\,\langle k| - 2\,|\kappa\rangle\,\langle\kappa| = 1 - 2\,|\kappa\rangle\,\langle\kappa|\,. \end{aligned} \tag{X.552}$$

7. Given the normalized states $|x\rangle$ and $|y\rangle$, with $\langle x\,|y\rangle = 0$; show that the operator $U = 2\,|x\rangle\,\langle x| - 1$ describes a reflection at $|x\rangle$ and $-U$ a reflection at $|y\rangle$.
Solution: We can represent an arbitrary state as $|z\rangle = a\,|x\rangle + b\,|y\rangle$. Then we have:

$$\begin{aligned} U\,|z\rangle &= 2\,|x\rangle\,\langle x\,|z\rangle - |z\rangle = 2a\,|x\rangle - a\,|x\rangle - b\,|y\rangle = a\,|x\rangle - b\,|y\rangle \\ -U\,|z\rangle &= -a\,|x\rangle + b\,|y\rangle\,. \end{aligned} \tag{X.553}$$

Hence, the operator U leaves the prefactor of $|x\rangle$ unchanged and modifies that of $|y\rangle$; accordingly, it is a reflection on $|x\rangle$. Analogously, one infers that $-U$ describes a reflection on $|y\rangle$; cf. Fig. X.10.

8. Given the normalized state

$$|\psi\rangle = \sum_{n=1}^{N} c_n\,|\varphi_n\rangle \;\text{ with }\; \langle\varphi_n\,|\varphi_m\rangle = \delta_{nm}. \tag{X.554}$$

The probability of measuring the state $|\varphi_k\rangle$ is thus given by $|c_k|^2$. We selectively amplify the amplitude $c_m \neq 0$ by the following unitary transformation U:

$$U : c_n \to \alpha c_n \text{ for } n \neq m;\;\; c_m \to \beta c_m \text{ for } n = m \tag{X.555}$$

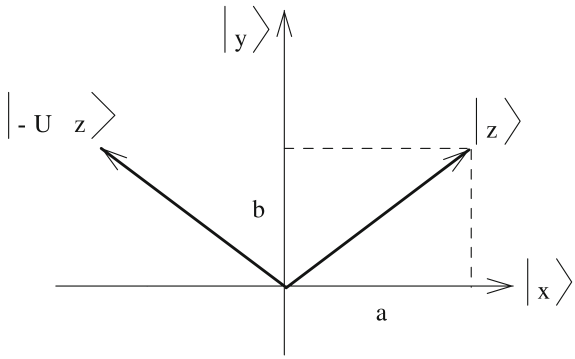
with suitably chosen α, β.

(a) How are α and β connected?
Solution: We have

$$\left|\psi'\right\rangle = U\,|\psi\rangle = \sum_{n=1}^{N} \left[\alpha + (\beta - \alpha)\,\delta_{nm}\right] c_n\,|\varphi_n\rangle\,, \tag{X.556}$$

and, due to the orthonormality of $\{|\varphi_n\rangle\}$, it follows that

Fig. X.10 The action of $-U = 1 - 2\,|x\rangle\langle x|$ on a general state



$$\langle\psi'\,|\psi'\rangle = 1 = \sum_{n=1}^{N} |\alpha + (\beta - \alpha)\,\delta_{nm}|^2\,|c_n|^2 = |\alpha|^2 + \left[|\beta|^2 - |\alpha|^2\right]|c_m|^2 . \tag{X.557}$$

This yields

$$1 = |\alpha|^2 + \left[|\beta|^2 - |\alpha|^2\right]|c_m|^2 \text{ or } |\beta|^2 = \frac{1 - |\alpha|^2}{|c_m|^2} + |\alpha|^2 ; \tag{X.558}$$

and for the probability, it follows

$$|c_n|^2 \to |\alpha|^2\,|c_n|^2 \quad \text{for } n \neq m;$$
$$|c_m|^2 \to |\beta|^2\,|c_m|^2 = 1 - |\alpha|^2 + |\alpha|^2\,|c_m|^2 \quad \text{for } n = m. \tag{X.559}$$

(b) How do the measurement probabilities behave under a k-fold iteration of U?
Solution: Clearly, a multiple application of the transformation gives

$$|c_n|^2 \to |\alpha|^2\,|c_n|^2 \to |\alpha|^4\,|c_n|^2 \to \cdots \tag{X.560}$$

and therefore

$$|c_m|^2 \to 1 - |\alpha|^2 + |\alpha|^2\,|c_m|^2 \to 1 - |\alpha|^4 + |\alpha|^4\,|c_m|^2 \to \cdots \tag{X.561}$$

With k iterations, we thus have:

$$|c_n|^2 \to |\alpha|^{2k}\,|c_n|^2 ; \quad |c_m|^2 \to 1 - |\alpha|^{2k} + |\alpha|^{2k}\,|c_m|^2 \to \cdots \tag{X.562}$$

(c) Specialize to the case of an initially uniform distribution $c_n = \frac{1}{\sqrt{N}}$ and $\alpha = \frac{1}{4}$. How often does one have to iterate in order to measure the state m with a probability of $w > 1 - 10^{-6}$ (assuming $N \gg 1$)?

Solution: With k iterations, we have

$$\begin{aligned}
&\frac{1}{N} \to \frac{1}{4^{2k}}\frac{1}{N} \text{ for } n \neq m; \\
&\frac{1}{N} \to 1 - \frac{1}{4^{2k}} + \frac{1}{4^{2k}}\frac{1}{N} = 1 - \frac{1}{4^{2k}}\frac{N-1}{N} \text{ for } n = m.
\end{aligned} \tag{X.563}$$

Hence, for $w > 1 - 10^{-6}$, it must hold that

$$1 - \frac{1}{4^{2k}}\frac{N-1}{N} > 1 - 10^{-6} \text{ or } 10^{-6} > \frac{1}{4^{2k}}\frac{N-1}{N} \approx \frac{1}{4^{2k}}, \tag{X.564}$$

and, due to $4^{2k} > 10^6$, it follows that

$$k > \frac{3}{\log 4} = 4.98\ldots \approx 5. \tag{X.565}$$

X.13 Exercises, Chap. 27

1. A system is in the polarization state $|r\rangle$. Using $w_P = tr\,(\rho P)$, calculate the probability of measuring the system in the state $|h\rangle$.
Solution: The density operator is $\rho = |r\rangle\langle r|$ and the projection operator $P = |h\rangle\langle h|$. It follows that:

$$\begin{aligned}
w_P &= \langle |h\rangle\langle h| \rangle = Sp\,(\rho\,|h\rangle\langle h|) = Sp\,(|r\rangle\langle r\,|h\rangle\langle h|) \\
&= \frac{1}{\sqrt{2}} Sp\,(|r\rangle\langle h|) = \frac{1}{\sqrt{2}}\left[\langle h\,|r\rangle\langle h\,|h\rangle + \langle v\,|r\rangle\langle h\,|v\rangle\right] = \frac{1}{2}.
\end{aligned} \tag{X.566}$$

2. A mixture is described by $\rho = \sum p_n |\varphi_n\rangle\langle\varphi_n|$, where $\{|\varphi_n\rangle\}$ is a CONS. Using $w_P = tr\,(\rho P)$, calculate the probability of measuring the system in the state $|\varphi_N\rangle$.
Solution: The projection operator is $P = |\varphi_N\rangle\langle\varphi_N|$. It follows that:

$$\begin{aligned}
w_P &= \langle P\rangle = Sp\,(\rho P) = Sp\,(\rho\,|\varphi_N\rangle\langle\varphi_N|) \\
&= \sum_n p_n Sp\,(|\varphi_n\rangle\langle\varphi_n|\,\varphi_N\rangle\langle\varphi_N|) = p_N Sp\,(|\varphi_N\rangle\langle\varphi_N|) = p_N.
\end{aligned} \tag{X.567}$$

3. The value function $V_{|\psi\rangle}$ is defined by $V_{|\psi\rangle}(F\,(A)) = F\left(V_{|\psi\rangle}(A)\right)$.

   (a) Prove for $[A, B] = 0$ the sum rule $V_{|\psi\rangle}(A + B) = V_{|\psi\rangle}(A) + V_{|\psi\rangle}(B)$.
   Solution: Due to $[A, B] = 0$, there exists an operator C such that $A = F(C)$ and $B = G(C)$. From this, it follows that $A + B = (F + G)\,(C)$, and thus

$$\begin{aligned} V_{|\psi\rangle}(A + B) &= V_{|\psi\rangle}((F + G)(C)) = (F + G)\, V_{|\psi\rangle}(C) \\ &= FV_{|\psi\rangle}(C) + GV_{|\psi\rangle}(C) = V_{|\psi\rangle}(F(C)) + V_{|\psi\rangle}(G(C)) \\ &= V_{|\psi\rangle}(A) + V_{|\psi\rangle}(B). \end{aligned} \tag{X.568}$$

(b) Prove for $[A, B] = 0$ the product rule $V_{|\psi\rangle}(A \cdot B) = V_{|\psi\rangle}(A) \cdot V_{|\psi\rangle}(B)$.
Solution: Due to the definition of the value function, we have:

$$V_{|\psi\rangle}(A^2) = V^2_{|\psi\rangle}(A) \text{ or } V_{|\psi\rangle}(A^n) = V^n_{|\psi\rangle}(A). \tag{X.569}$$

We can again assume $A = F(C)$ and $B = G(C)$. Expanding the functions in power series, we get

$$\begin{aligned} V_{|\psi\rangle}(A \cdot B) &= V_{|\psi\rangle}(F(C) \cdot G(C)) \\ &= V_{|\psi\rangle}(\sum_n F_n C^n \cdot \sum_m G_m C^m) = V_{|\psi\rangle}(\sum_{n,m} F_n G_m C^{n+m}) \\ &= \sum_{n,m} F_n G_m V_{|\psi\rangle}\left(C^{n+m}\right) = \sum_{n,m} F_n G_m V^{n+m}_{|\psi\rangle}(C) \\ &= \sum_{n,m} F_n V^n_{|\psi\rangle}(C)\, G_m V^m_{|\psi\rangle}(C) = \sum_n F_n V^n_{|\psi\rangle}(C) \cdot \sum_m G_m V^m_{|\psi\rangle}(C) \\ &= \sum_n F_n V_{|\psi\rangle}\left(C^n\right) \cdot \sum_m G_m V_{|\psi\rangle}\left(C^m\right) \\ &= V_{|\psi\rangle}\left(\sum_n F_n C^n\right) \cdot V^m_{|\psi\rangle}\left(\sum_m G_m C^m\right) \\ &= V_{|\psi\rangle}(A) \cdot V^m_{|\psi\rangle}(B). \end{aligned} \tag{X.570}$$

(c) Show that $V_{|\psi\rangle}(1) = 1$.
Solution: Let 1 be the unit operator. Due to the product rule, it holds that $V_{|\psi\rangle}(B) = V_{|\psi\rangle}(1 \cdot B) = V_{|\psi\rangle}(1) \cdot V_{|\psi\rangle}(B)$, and it follows that $V_{|\psi\rangle}(1) = 1$, where it is supposed that there is at least one quantity B for which $V_{|\psi\rangle}(B) \neq 0$.

4. Given the polarization operators $P_L, P_{L'}$ and P_C (or the corresponding Pauli matrices, see (27.11)):

(a) Determine (once more) their eigenvalues and eigenvectors.
Solution: Due to $P_A^2 = 1$, the eigenvalues are given by $\lambda = \pm 1$. For e.g. $P_{L'}$, it holds that

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \pm \begin{pmatrix} a \\ b \end{pmatrix}, \tag{X.571}$$

and this gives the normalized eigenvectors:

$$P_{L'}: \left|h'\right\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \text{ for } \lambda_{L'} = 1; \; \left|v'\right\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix} \text{ for } \lambda_{L'} = -1. \quad \text{(X.572)}$$

Remark: As always, the vectors are defined up to a phase; we choose it in such a way that $\left|h'\right\rangle$ and $\left|v'\right\rangle$ arise from $|h\rangle$ and $|v\rangle$ by an active rotation. Analogously, it follows that

$$\begin{aligned} P_C: |r\rangle &= \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix} \text{ for } \lambda_C = 1; \; |l\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix} \text{ for } \lambda_Z = -1 \\ P_L: |h\rangle &= \begin{pmatrix}1\\0\end{pmatrix} \text{ for } \lambda_L = 1; \; |v\rangle = \begin{pmatrix}0\\1\end{pmatrix} \text{ for } \lambda_L = -1. \end{aligned} \quad \text{(X.573)}$$

(b) Express the eigenvectors of P_C and $P_{L'}$ in terms of those of P_L.
Solution: We have

$$\begin{aligned} \left|h'\right\rangle &= \frac{|h\rangle + |v\rangle}{\sqrt{2}}; \; \left|v'\right\rangle = \frac{-|h\rangle + |v\rangle}{\sqrt{2}} \\ |r\rangle &= \frac{|h\rangle + i\,|v\rangle}{\sqrt{2}}; \; |l\rangle = \frac{|h\rangle - i\,|v\rangle}{\sqrt{2}}. \end{aligned} \quad \text{(X.574)}$$

The inversion reads:

$$\begin{aligned} |h\rangle &= \frac{\left|h'\right\rangle - \left|v'\right\rangle}{\sqrt{2}}; \; |v\rangle = \frac{\left|h'\right\rangle + \left|v'\right\rangle}{\sqrt{2}} \\ |h\rangle &= \frac{|r\rangle + |l\rangle}{\sqrt{2}}; \; |v\rangle = \frac{|r\rangle - |l\rangle}{\sqrt{2}i}. \end{aligned} \quad \text{(X.575)}$$

5. Given the GHZ state

$$|\psi\rangle_{\pm} = \frac{|h, h, h\rangle \pm |v, v, v\rangle}{\sqrt{2}} \quad \text{(X.576)}$$

corresponding to an LLL measurement; rewrite this for a CCL' measurement (plus $CL'C$ and $L'CC$) (27.12) and for an $L'L'L'$ measurement (27.14).

(a) Solution for CCL': With the results from the previous exercise, we have

$$\begin{aligned} \sqrt{2}\,|\psi\rangle_{\pm} = &\left(\frac{|r\rangle + |l\rangle}{\sqrt{2}}\right)_1 \left(\frac{|r\rangle + |l\rangle}{\sqrt{2}}\right)_2 \left(\frac{\left|h'\right\rangle - \left|v'\right\rangle}{\sqrt{2}}\right)_3 \\ &\pm \left(\frac{|r\rangle - |r\rangle}{\sqrt{2}i}\right)_1 \left(\frac{|r\rangle - |l\rangle}{\sqrt{2}i}\right)_2 \left(\frac{\left|h'\right\rangle + \left|v'\right\rangle}{\sqrt{2}}\right)_3. \end{aligned} \quad \text{(X.577)}$$

Expanding the first two ratios yields:

$$\sqrt{2}\,|\psi\rangle_{\pm} = \frac{|r,r\rangle + |r,l\rangle + |l,r\rangle + |l,l\rangle}{2}\left(\frac{|h'\rangle - |v'\rangle}{\sqrt{2}}\right)_3 \mp \frac{|r,r\rangle - |r,l\rangle - |l,r\rangle + |l,l\rangle}{2}\left(\frac{|h'\rangle + |v'\rangle}{\sqrt{2}}\right)_3. \quad (X.578)$$

From this, it follows that

$$|\psi\rangle_+ = \frac{|r,l,h'\rangle + |l,r,h'\rangle - |r,r,v'\rangle - |l,l,v'\rangle}{2}$$
$$|\psi\rangle_- = \frac{|r,r,h'\rangle + |l,l,h'\rangle - |r,l,v'\rangle - |l,r,v'\rangle}{2}. \quad (X.579)$$

The results for $CL'C$ and $L'CC$ follow by cyclic permutation. It is clear that if we have two readings, we can predict the third with certainty; the combination $|r,r,?\rangle$ can be only $|r,r,v'\rangle$ for $|\psi\rangle_+$, $|l,?,h'\rangle$ only $|l,r,h'\rangle$, etc.

(b) Solution for $L'L'L'$: We have initially:

$$\sqrt{2}\,|\psi\rangle_{\pm} = \left(\frac{|h'\rangle - |v'\rangle}{\sqrt{2}}\right)_1\left(\frac{|h'\rangle - |v'\rangle}{\sqrt{2}}\right)_2\left(\frac{|h'\rangle - |v'\rangle}{\sqrt{2}}\right)_3 \pm \left(\frac{|h'\rangle + |v'\rangle}{\sqrt{2}}\right)_1\left(\frac{|h'\rangle + |v'\rangle}{\sqrt{2}}\right)_2\left(\frac{|h'\rangle + |v'\rangle}{\sqrt{2}}\right)_3. \quad (X.580)$$

We expand again the first two ratios:

$$\sqrt{2}\,|\psi\rangle_{\pm} = \frac{|h',h'\rangle - |v',h'\rangle - |h',v'\rangle + |v',v'\rangle}{2}\left(\frac{|h'\rangle - |v'\rangle}{\sqrt{2}}\right)_3 \pm \frac{|h',h'\rangle + |v',h'\rangle + |h',v'\rangle + |v',v'\rangle}{2}\left(\frac{|h'\rangle + |v'\rangle}{\sqrt{2}}\right)_3, \quad (X.581)$$

and obtain

$$|\psi\rangle_+ = \frac{|h',h',h'\rangle + |v',v',h'\rangle + |v',h',v'\rangle + |h',v',v'\rangle}{2}$$
$$|\psi\rangle_- = -\frac{|v',h',h'\rangle + |h',v',h'\rangle + |h',h',v'\rangle + |v',v',v'\rangle}{2}. \quad (X.582)$$

Two readings again with certainty determine the third one. All in all we have an odd number of states $|h'\rangle$ for $|\psi\rangle_+$, and an even number for $|\psi\rangle_-$.

6. The following combinations of the polarization operators (27.10) are given:

$$Q_1 = P_{1L'}P_{2C}P_{3C};\ Q_2 = P_{1C}P_{2L'}P_{3C}$$
$$Q_3 = P_{1C}P_{2C}P_{3L'};\ Q = P_{1L'}P_{2L'}P_{3L'}. \tag{X.583}$$

The numerical index denotes the space in which the particular polarization operator acts. We use in the following the fact that operators from different spaces commute, e.g. $P_{1L'}P_{2C} = P_{2C}P_{1L'}$. In addition, we have $P_{nL'}P_{nC} = -P_{nC}P_{nL'}$ as well as $P_{nC}^2 = P_{nL'}^2 = 1$.

(a) Show that the three operators Q_i have the eigenvalues ± 1.
Solution: Evidently, we have

$$Q_i^2 = 1, \tag{X.584}$$

and the proposition follows immediately from this.

(b) Show that the three operators Q_i commute pairwise.
Solution: We have e.g.

$$\begin{aligned} Q_1 Q_2 &= P_{1L'}P_{2C}P_{3C}P_{1C}P_{2L'}P_{3C} = P_{1L'}P_{2C}P_{1C}P_{2L'} \\ &= -P_{1C}P_{2C}P_{1L'}P_{2L'} = P_{1C}P_{2L'}P_{1L'}P_{2C} \\ &= P_{1C}P_{2L'}P_{3C}P_{1L'}P_{2C}P_{3C} = Q_2 Q_1. \end{aligned} \tag{X.585}$$

(c) Show that the states

$$|\psi\rangle_\pm = \frac{|h, h, h\rangle \pm |v, v, v\rangle}{\sqrt{2}} \tag{X.586}$$

are common eigenstates of the three operators Q_i with the eigenvalues ∓ 1, as well as eigenstates of the operator Q with the eigenvalues ± 1.
Solution: With

$$\begin{aligned} P_{L'}|h\rangle = |v\rangle;\ P_{L'}|v\rangle = |h\rangle \\ P_C|h\rangle = i|v\rangle;\ P_C|v\rangle = -i|h\rangle, \end{aligned} \tag{X.587}$$

it follows that e.g.

$$\begin{aligned} Q_1|\psi\rangle_\pm &= P_{1L'}P_{2C}P_{3C}\frac{|h, h, h\rangle \pm |v, v, v\rangle}{\sqrt{2}} \\ &= \frac{i^2|v, v, v\rangle \pm (-i)^2|h, h, h\rangle}{\sqrt{2}} = \mp|\psi\rangle_\pm, \end{aligned} \tag{X.588}$$

and analogously for Q_2 and Q_3. For Q, we have:

$$\begin{aligned} Q|\psi\rangle_\pm &= P_{1L'}P_{2L'}P_{3L'}\frac{|h, h, h\rangle \pm |v, v, v\rangle}{\sqrt{2}} \\ &= \frac{|v, v, v\rangle \pm |h, h, h\rangle}{\sqrt{2}} = \pm|\psi\rangle_\pm. \end{aligned} \tag{X.589}$$

Further Reading

1. J. Audretsch (ed.), *Verschränkte Welt-Faszination der Quanten (Entangled World-Fascination of Quantum, in German)* (Wiley-VCH, Weinheim, 2002)
2. J. Audretsch, *Entangled Systems* (Wiley-VCH, Weinheim, 2007)
3. J.-L. Basdevant, J. Dalibard, *Quantum Mechanics* (Springer, Berlin, 2002)
4. D.R. Bes, *Quantum Mechanics* (Springer, Berlin, 2004)
5. B.H. Bransden, C.J. Joachain, *Quantum Mechanics* (Pearson Education Limited, Harlow, 2000)
6. C. Cohen-Tannoudji, B. Diu, F. Laloë, *Quantum Mechanics*, vols. 1 & 2 (Hermann Paris/Wiley, New York, 1977)
7. F. Embacher, Homepage with much material about quantum theory (also for school), University Vienna (2012), http://homepage.univie.ac.at/franz.embacher/
8. R.P. Feynman, R.B. Leighton, M. Sand, *Quantum Mechanics*. The Feynman Lectures on Physics, vol. 3 (Addison-Wesley Reading, Massachusetts, 1965)
9. T. Fließbach, *Quantenmechanik (Quantum Mechanics, in German)*, 3rd edn. (Spektrum Akademischer Verlag, Heidelberg, 2000)
10. K. Gottfried, T.-M. Yan, *Quantum Mechanics: Fundamentals* (Springer, New York, 2006)
11. K.T. Hecht, *Quantum Mechanics* (Springer, New York, 2000)
12. C.J. Isham, *Quantum Theory-Mathematical and Structural Foundations* (Imperial College Press, London, 2008)
13. T. Lancaster, S.J. Blundell, *Quantum Field Theory for the Gifted Amateur* (Oxford University Press, Oxford, 2014)
14. R.D. Klauber, *Student Friendly Quantum Field Theory*, 2nd edn. (Sandrove Press, Fairfield, 2015)
15. M. Le Bellac, *Quantum Physics* (Cambridge University Press, Cambridge, 2006)
16. H. Lüth, *Quantenphysik in der Nanowelt (Quantum Physics in the Nanoworld, in German)* (Springer, Berlin, 2009)
17. E. Merzbacher, *Quantum Mechanics*, 3rd edn. (Wiley, New York, 1998)
18. A. Messiah, *Quantum Mechanics*, vols. 1 & 2 (North-Holland Publishing Company, Amsterdam, 1964)
19. Münchener Internetprojekt zur Lehrerfortbildung in Quantenmechanik (Munich internet project for teacher training in quantum mechanics, in German) (2012), http://homepages.physik.uni-muenchen.de/milq/
20. G. Münster, *Quantentheorie (Quantum Theory, in German)* (Walter de Gruyter, Berlin, 2006)
21. W. Nolting, *Grundkurs Theoretische Physik 5, Quantenmechanik, Teil 1: Grundlagen und Quantenmechanik Teil 2: Methoden und Anwendungen (Quantum Mechanics, Part 1: Basics*

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

and Part 2: Methods and Applications, in German) (Verlag Zimmermann-Neufang, Ulmen, 1992)
22. A. Peres, *Quantum Theory-Concepts and Methods* (Kluwer Academic Publishers, Doordrecht, 1995)
23. A.I.M. Rae, *Quantum Mechanics*, 5th edn. (Taylor and Francis, New York, 2008)
24. H. Rollnik, *Quantentheorie 1 & 2 (Quantum Theory 1 & 2, in German)*, 2nd edn. (Springer, Berlin, 2003)
25. H. Schulz, *Physik mit Bleistift (Physics with a Pencil, in German)*, 4th edn. (Verlag Harri Deutsch, Frankfurt am Main, 2001)
26. F. Schwabl, *Quantum Mechanics*, 3rd edn. (Springer, Berlin, 2002)
27. N. Zettili, *Quantum Mechanics, Concepts and Applications*, 2nd edn. (Wiley, New York, 2009)

Index of Volume 1

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5

Index of Volume 2

J. Pade, *Quantum Mechanics for Pedestrians 2*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00467-5