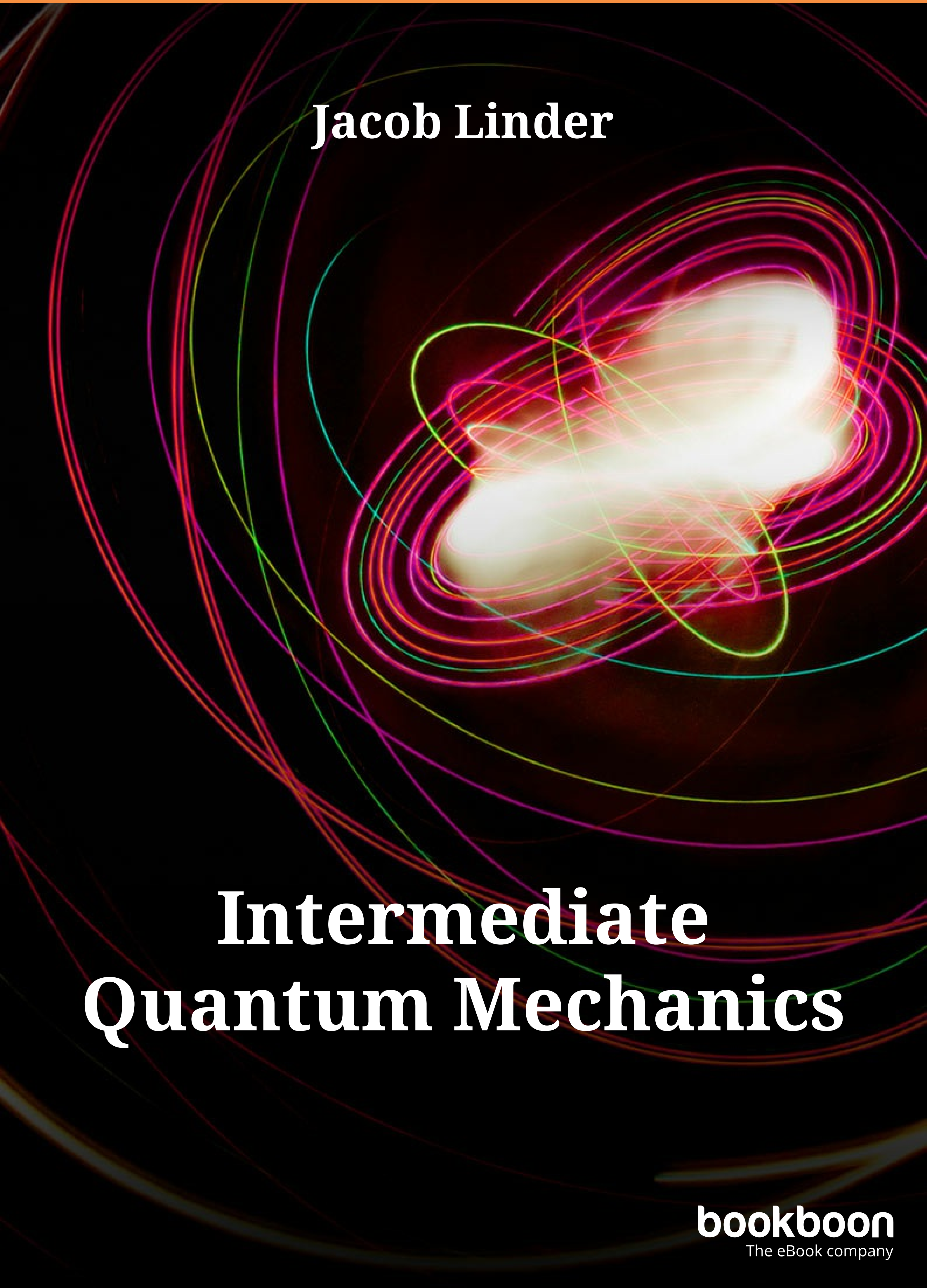


Jacob Linder

Intermediate Quantum Mechanics

bookboon
The eBook company

JACOB LINDER

INTERMEDIATE QUANTUM MECHANICS

Intermediate Quantum Mechanics
1st edition

ISBN 978-87-403-1783-1
Peer review by Dr. Takehito Yokoyama, Tokyo Institute of Technology, Japan

CONTENTS

Preface

The aim of this book is to present fundamental concepts in quantum mechanics and a general mathematical formalism beyond the wavefunction framework taught in introductory quantum mechanics courses. This includes topics such as Dirac formalism with bra- and ket-vectors in Hilbert space, Heisenberg formalism with matrices, approximation methods in quantum mechanics, scattering theory, atoms and electrons in magnetic fields, coherent states, field quantization and radiation theory, and the density matrix formalism. In addition to explaining the underlying theory in a detailed manner, we shall also provide a number of examples that will illustrate the formalisms "in action".

This book is primarily based on my lecture notes from teaching this class to undergraduate students, and the notes in turn are based on the book "Kvantemekanikk" by P. C. Hemmer. I have also included additional topics and instructive examples which hopefully will allow the reader to obtain a more thorough physical understanding of the material. This book is suitable as material for a full-semester course in intermediate quantum mechanics at the undergraduate level.

It is my goal that students who study this book afterwards will find themselves well prepared to dig deeper into the remarkable world of theoretical physics at a more advanced level. I welcome feedback on the book (including any typos that you may find, although I have endeavored to eliminate as many of them as possible) and hope that you will have an exciting time reading it!

Jacob Linder (jacob.linder@ntnu.no)
Norwegian University of Science and Technology
Trondheim, Norway

About the author

J.L. holds a position as Professor of Physics at the Norwegian University of Science and Technology. His research is focused on theoretical quantum condensed matter physics and he has received several prizes for his Ph.D work on the interplay between superconductivity and magnetism. He has also received the American Physical Society "Outstanding Referee" award, selected among over 60.000 active referees. In teaching courses such as Quantum Mechanics, Classical Mechanics, and Particle Physics for both undergraduate and graduate students, he has invariably received high scores from the students for his pedagogical qualities and lectures. His webpage is found here. He has also written the book "Introduction to Lagrangian & Hamiltonian Mechanics" together with Prof. Iver Brevik, which is freely available on Bookboon.

I. GENERAL FORMULATION OF QUANTUM MECHANICS

Learning goals. After reading this chapter, the student should:

- Be able to understand and use Dirac's bra-ket notation.
- Know the fundamental axioms of the general formulation of QM.
- Know how the general formulation and the wavemechanics formulation of QM are related.

Introductory quantum mechanics (QM) utilizes a position-representation where one works with wavefunctions $\psi = \psi(\boldsymbol{r})$. However, this is in fact just a special case of a more general theory. The general theory is important because some QM systems cannot be treated by wavefunctions in position space, such as the spin degree of freedom. We therefore develop the foundation for the general theory in what follows.

A. Dirac's bra-ket notation

We introduce a new formulation where a QM state is described by a state vector $|\psi\rangle$ in a complex linear vector space $\mathscr{H}$, namely the so-called Hilbert space. The Hilbert space $\mathscr{H}$ may have a finite or infinite dimension, and in often cases the latter. For instance, we need infinite Hilbert spaces to represent a vector describing continuous variables (such as position). In contrast, only a two-dimensional Hilbert space is required to describe a single spin-1/2 state. We will show this explicitly later on. For now, you may simply think of $\mathscr{H}$ as the space where the state vector $|\psi\rangle$ resides. Mathematically, $\mathscr{H}$ is required in order to perform operations such as inner products between state vectors in a well-defined manner.

There are different notations which are used for the state vector. A common convention is to denote the state with its quantum numbers. For instance, stationary states in a Coulomb-field would then be written as $|nlm\rangle$, where $\{n, l, m\}$ are the quantum numbers characterizing the eigenstates of the system (as treated in introductory courses to quantum mechanics). Generally, the state vector may also depend on time. In what follows, we usually suppress the t-dependence notation-wise unless it is of importance.

For any state $|a\rangle$ in $\mathscr{H}$, there is assigned a dual vector $\langle a|$ in a dual vector space. The relation between the two state vectors is that the scalar product

$$\langle a| \cdot |b\rangle \equiv \langle a|b\rangle \tag{1.1}$$

is defined as a complex number with the property

$$\langle a|b\rangle = \langle b|a\rangle^*. \tag{1.2}$$

The two states $|a\rangle$ and $|b\rangle$ are orthogonal if their inner product $\langle a|b\rangle = 0$. The notation used here is due to Dirac and known as *bra-ket* notation:

$$\langle \ldots | \equiv \text{bra}, \;\; | \ldots \rangle \equiv \text{ket}. \tag{1.3}$$

If a vector is multiplied with a number c, the corresponding dual vector must be multiplied by c^*. To see this, let $|a'\rangle = c|a\rangle$. It follows that $\langle b|a'\rangle = c\langle b|a\rangle$ and thus

$$\langle a'|b\rangle = \langle b|a'\rangle^* = c^*\langle b|a\rangle^* = c^*\langle a|b\rangle. \tag{1.4}$$

It is then clear that $\langle a'| = c^*\langle a|$.

In a n-dimensional vector space, we may choose n linearly independent vectors $|1\rangle, |2\rangle, \ldots |n\rangle$ as basis vectors and expand an arbitrary state vector $|\psi\rangle$ in these:

$$|\psi\rangle = \sum_{k=1}^{n} c_k |k\rangle, \tag{1.5}$$

where c_k are complex numbers. Assume for simplicity that these basis vectors are orthonormal, so that $\langle k|m\rangle = \delta_{km}$. We allow the dimension n to not necessarily be finite. It follows that $c_m = \langle m|\psi\rangle$, so that we may write

$$|\psi\rangle = \sum_k \langle k|\psi\rangle |k\rangle. \tag{1.6}$$

In turn, this can be written as

$$|\psi\rangle = \sum_k |k\rangle\langle k| \cdot |\psi\rangle \tag{1.7}$$

(we simply interchanged the position of $\langle k|\psi\rangle$ and $|k\rangle$ which is fine since $\langle k|\psi\rangle$ is a scalar) which means that we must have

$$\sum_k |k\rangle\langle k| = 1. \tag{1.8}$$

This is the so-called *completeness relation* which will turn out to be very useful. The corresponding relation for usual vectors in three dimensional Cartesian space can in fact be written in a similar fashion:

$$\sum_k (\boldsymbol{e}_k)\boldsymbol{e}_k \cdot = (\boldsymbol{e}_x)\boldsymbol{e}_x \cdot + (\boldsymbol{e}_y)\boldsymbol{e}_y \cdot + (\boldsymbol{e}_z)\boldsymbol{e}_z \cdot = 1, \tag{1.9}$$

because using this operator on a vector $\boldsymbol{A}$ is equivalent to the identity operation:

$$\boldsymbol{e}_x A_x + \boldsymbol{e}_y A_y + \boldsymbol{e}_z A_z = \boldsymbol{A}. \tag{1.10}$$

While $\langle a|b\rangle$ is the inner product and equal to a complex number in general, the outer product of the vectors $|a\rangle$ and $|b\rangle$ is $|a\rangle\langle b|$ and is generally equal to an operator. For instance, $|k\rangle\langle k|$ is a projection operator that projects a state vector onto the $|k\rangle$-axis.

Some basis vector sets $\{|k\rangle\}$ are such that k takes on continuous values. Then, we replace the summation with an integration and also a delta function normalization:

$$\langle k|k'\rangle = \delta(k - k'). \tag{1.11}$$

The expansion of a state vector $|\psi\rangle$ using such basis vectors then takes the form

$$|\psi\rangle = \int dk\; c(k)|k\rangle. \tag{1.12}$$

Multiplying from the left with $\langle k'|$, we find the expansion coefficients

$$\langle k'|\psi\rangle = \int dk\; c(k)\langle k'|k\rangle = \int dk\; c(k)\delta(k'-k) = c(k'). \tag{1.13}$$

We may thus write that

$$|\psi\rangle = \int dk \langle k|\psi\rangle|k\rangle. \tag{1.14}$$

The completeness relation for continuous variables then takes the form

$$\int dk|k\rangle\langle k| = 1. \tag{1.15}$$

The norm $||f||$ of a vector $|f\rangle$ is defined as

$$||f|| = \sqrt{\langle f|f\rangle} \geq 0. \tag{1.16}$$

We can see that $||f||$ is always real and non-negative by using the completeness relation we derived. It follows from the property:

$$\langle f|f\rangle = \sum_k \langle f|k\rangle\langle k|f\rangle = \sum_k |\langle k|f\rangle|^2 \tag{1.17}$$

since $\langle f|k\rangle = \langle k|f\rangle^*$.

B. Operators and eigenvectors

An operator in Hilbert space $\mathscr{H}$ is an image of $\mathscr{H}$ on itself. This means that the operator A assigns a vector $|c\rangle$ to any vector $|a\rangle$ according to:

$$A|a\rangle = |c\rangle. \tag{1.18}$$

The adjoint operator $A^\dagger$ is defined by

$$\langle a|A^\dagger|b\rangle = \langle b|A|a\rangle^* \tag{1.19}$$

which must hold for any two vectors $|a\rangle$ and $|b\rangle$ in $\mathscr{H}$. By setting $A|a\rangle = |c\rangle$, we may write Eq. (1.19) as

$$\langle a|A^\dagger|b\rangle = \langle b|c\rangle^* = \langle c|b\rangle. \tag{1.20}$$

It then follows that $\langle c| = \langle a|A^\dagger$. We have thus shown that the dual vector of $A|a\rangle$ is $\langle a|A^\dagger$. The following properties of the adjoint operation follow from our definitions so far (try to prove them yourself!)

- $(A^\dagger)^\dagger = A$
- $(\alpha A)^\dagger = \alpha^* A^\dagger$ where α is a constant
- $(AB)^\dagger = B^\dagger A^\dagger$.

An operator is self-adjoint (also known as Hermitian) if

$$A^\dagger = A. \tag{1.21}$$

It follows that for such operators

$$\langle a|A|a\rangle = \langle a|A|a\rangle^* \to \langle a|A|a\rangle \in \Re. \tag{1.22}$$

We define an eigenvector of A to be $|\alpha\rangle$ where

$$A|\alpha\rangle = \lambda_\alpha|\alpha\rangle. \tag{1.23}$$

The number λ_α is the eigenvalue. The collection of eigenvalues for the operator A are known as the spectrum of A. An important observation is that:

$$\boxed{\text{Eigenvalues of Hermitian operators } A \text{ are real.}}$$

This follows since $\lambda_\alpha = \frac{\langle\alpha|A|\alpha\rangle}{\langle\alpha|\alpha\rangle} \in \Re$, as shown above. Representing a physical observable with a Hermitian operator thus guarantees that the corresponding eigenvalues are real, as they should be for a measurable quantity. The set of eigenvectors $\{|\alpha\rangle\}$ for an operator corresponding to a physical quantity is assumed to be a complete set. This means that such eigenvectors may be used as basis vectors.

C. The axioms of the general formulation of QM

The general formulation of QM, which we have now established the notation for, is based on the following postulates.

A: To any observable quantity F, one assigns a linear Hermitian operator $\hat{F}$ in Hilbert space. The operators of a generalized coordinate q_n and the corresponding generalized momentum p_n satsify the commutation relation $[\hat{q}_n, \hat{p}_n] = \mathrm{i}\hbar$.

B: The state of a physical system is described by a state vector $|\psi(t)\rangle$ in a Hilbert space. It has the property $\langle\psi(t)|\psi(t)\rangle = 1$ and satisfies the time-dependent Schrödinger-equation

$$\mathrm{i}\hbar\partial_t|\psi(t)\rangle = \hat{H}|\psi(t)\rangle. \tag{1.24}$$

Here, $\hat{H}$ is the Hamilton operator.

C: The expectation value of an observable quantity F in the state $|\psi\rangle$ is $\langle F\rangle = \langle\psi|\hat{F}|\psi\rangle$.

D: The measurement of an observable quantity F yields as a result one of the eigenvalues f_n of the operator $\hat{F}$. An observable quantity is defined as a property of the system's state which may be determined by performing physical operations on the system (such as subjecting a charged particle to a magnetic field and reading off its position).

With these postulates, we can now describe QM with different sets of basis vectors. To begin with, we will look at this in more detail using the position representation and establish how this is related to the wavefunction formalism in introductory courses of quantum mechanics.

D. Different representations

We start with the position representation and consider motion only in 1D, in order to keep the notation simple. The eigenvectors of the position operator $\hat{x}$ are dneoted $|x'\rangle$ where x' is the eigenvalue:

$$\hat{x}|x'\rangle = x'|x'\rangle. \tag{1.25}$$

We then have $x' \in (-\infty, \infty)$. The total state vector may be expanded as

$$|\psi\rangle = \int dx' \langle x'|\psi\rangle|x'\rangle. \tag{1.26}$$

The complex number $\langle x'|\psi\rangle$ is the contribution to the state vector $|\psi\rangle$ from position x'. Hence, this is in fact nothing but the familiar wavefunction in position space:

$$\psi(x') = \langle x'|\psi\rangle. \tag{1.27}$$

We see that $\psi(x) = \langle x|\psi\rangle$ are the components of $|\psi\rangle$ with the basis vectors $|x\rangle$. If x instead took discrete values, we could have written

$$|\psi\rangle = \begin{pmatrix} \psi(x_1) \\ \psi(x_2) \\ \dots \end{pmatrix}. \tag{1.28}$$

Since x is a continuous variable, we must δ-function normalize the basis-vectors:

$$\langle x''|x'\rangle = \delta(x'' - x'). \tag{1.29}$$

The scalar product between $|\psi_1\rangle$ and $|\psi_2\rangle$ may be written as

$$\langle\psi_1|\psi_2\rangle = \int dx \langle\psi_1|x\rangle\langle x|\psi_2\rangle \tag{1.30}$$

$$= \int dx \langle x|\psi_1\rangle^* \langle x|\psi_2\rangle \tag{1.31}$$

$$= \int dx\ \psi_1^*(x)\psi_2(x), \tag{1.32}$$

where we made use of the completeness relation $\int dx|x\rangle\langle x| = 1$. Let us also consider how to work with operators in this representation. The expectation value of $\hat{F}$ may be written as:

$$\langle\psi|\hat{F}|\psi\rangle = \int dx'' \int dx' \langle\psi|x''\rangle\langle x''|\hat{F}|x'\rangle\langle x'|\psi\rangle. \tag{1.33}$$

The first and last factors inside the integral are wavefunctions, as we showed previously, so it remains to clarify what the *matrix elements* $\langle x''|\hat{F}|x'\rangle$ are. If $\hat{F} = \hat{x}$, it is simple. We then have:

$$\langle x''|\hat{x}|x'\rangle = x'\langle x''|x'\rangle = x'\delta(x'' - x'). \tag{1.34}$$

More generally, if $\hat{F}$ is a function of $\hat{x}$ [$\hat{F} = F(\hat{x})$], then

$$\langle x''|F(\hat{x})|x'\rangle = F(x')\delta(x'' - x'). \tag{1.35}$$

This follows for any power of $\hat{x}$ since $\hat{x}^n|x'\rangle = (x')^n|x'\rangle$, and thus the same is true for any function $F(\hat{x})$ that may be expanded in powers of $\hat{x}$.

What about the case $\hat{F} = \hat{p}_x \equiv \hat{p}$? We know that $[\hat{x}, \hat{p}] = \mathrm{i}\hbar$, and thus

$$\langle x''|\hat{x}\hat{p} - \hat{p}\hat{x}|x'\rangle = \mathrm{i}\hbar\langle x''|x'\rangle = \mathrm{i}\hbar\delta(x'' - x'). \tag{1.36}$$

The second term on the l.h.s. is

$$\langle x''| - \hat{p}\hat{x}|x'\rangle = -x'\langle x''|\hat{p}|x'\rangle. \tag{1.37}$$

The first term on the l.h.s. may be computed as follows:

$$\begin{aligned}\langle x''|\hat{x}\hat{p}|x'\rangle &= \int dx_1 \langle x''|\hat{x}|x_1\rangle\langle x_1|\hat{p}|x'\rangle \\ &= \int dx_1 x_1 \delta(x'' - x_1)\langle x_1|\hat{p}|x'\rangle \\ &= x''\langle x''|\hat{p}|x'\rangle. \end{aligned}\tag{1.38}$$

An alternative way to show this explicitly for the second term is as follows. First, note that if $\hat{A}|a\rangle = \lambda|a\rangle$, then $\langle a|\hat{A}^\dagger = \langle a|\lambda^*$. We then see that

$$\langle x''|\hat{p}\hat{x}|x'\rangle = \langle x'|(\hat{p}\hat{x})^\dagger|x''\rangle^* = \langle x'|\hat{x}^\dagger\hat{p}^\dagger|x''\rangle^*. \tag{1.39}$$

But $\hat{x}$ and $\hat{p}$ must be Hermitian so that $\hat{x}^\dagger = \hat{x}$ and similarly for $\hat{p}$. By using this, we obtain that

$$\begin{aligned}\langle x'|\hat{x}^\dagger\hat{p}^\dagger|x''\rangle^* &= x'\langle x'|\hat{p}|x''\rangle^* \\ &= x'\langle x''|\hat{p}|x'\rangle, \end{aligned}\tag{1.40}$$

which is consistent with Eq. (1.37). Combining the results we obtained so far, we then have that

$$(x'' - x')\langle x''|\hat{p}|x'\rangle = \mathrm{i}\hbar\delta(x'' - x'). \tag{1.41}$$

Now, a fundamental property of the δ-function is that $x\frac{\partial\delta(x)}{\partial x} = -\delta(x)$. We now use this by letting $x \equiv x'' - x'$ and hold x' to be constant. It follows that

$$(x'' - x')\frac{\partial}{\partial x''}\delta(x'' - x') = -\delta(x'' - x'). \tag{1.42}$$

We can then rewrite Eq. (1.41) to

$$\langle x''|\hat{p}|x'\rangle = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x''}\delta(x'' - x'). \tag{1.43}$$

This can be further generalized to a power $\hat{p}^n$:

$$\langle x''|\hat{p}^n|x'\rangle = \Big(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x''}\Big)^n\delta(x'' - x'). \tag{1.44}$$

Since we have now proven that for an arbitrary function $F(p)$, we have

$$\langle x''|F(\hat{p})|x'\rangle = F\Big(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x''}\Big)\delta(x'' - x'), \tag{1.45}$$

it follows that in the most general case where the operator depends on *both* $\hat{p}$ and $\hat{x}$, we have:

$$\langle x''|F(\hat{p},\hat{x})|x'\rangle = F\Big(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x''}, x''\Big)\delta(x'' - x'), \tag{1.46}$$

Since we now know this expectation value, we can finally go back and evaluate the expression we started out with:

$$\begin{aligned}\langle\psi|\hat{F}|\psi\rangle &= \int dx'' \int dx' \langle\psi|x''\rangle\langle x''|\hat{F}|x'\rangle\langle x'|\psi\rangle \\ &= \int dx'' \int dx' \psi^*(x'')F\Big(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x''}, x''\Big)\delta(x'' - x')\psi(x') \\ &= \int dx' \psi^*(x')F\Big(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x'}, x'\Big)\psi(x'). \end{aligned}\tag{1.47}$$

In the end, we see that this is precisely how we are used to evaluate expectation values in the wavefunction formulation. Hence, there is consistency between the general formulation of QM and the position representation.

We may also show that the two versions of the Schrödinger equation (SE) are consistent. The general formula is:

$$\mathrm{i}\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle \tag{1.48}$$

and can be brought to the position representation by multiplying with $\langle x|$ from the left side, so that one obtains

$$\mathrm{i}\hbar\partial_t\langle x|\psi\rangle = \langle x|\hat{H}|\psi\rangle = \int dx'\langle x|\hat{H}|x'\rangle\langle x'|\psi\rangle. \tag{1.49}$$

We previously established that $\langle x|\psi\rangle = \psi(x)$. For a Hamiltonian operator $\hat{H} = H(\hat{p},\hat{x})$, it follows that

$$\mathrm{i}\hbar\partial_t\psi(x) = \int dx' H\Big(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}, x\Big)\delta(x-x')\psi(x') = H\Big(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}, x\Big)\psi(x). \tag{1.50}$$

Summarizing, we see that when the eigenvectors for the position operator are used as basis vectors, the general formulation of QM is reduced to wavemechanics in position space. In the same manner as above, the wavemechanics in momentum space is contained in the general formulation of QM. In this case, we want to use the eigenvectors $|p\rangle$ of the momentum operator $\hat{p}$ as basis vectors. The wavefunction in the momentum representation is then $\phi(p) = \langle p|\psi\rangle$.

Interestingly, the wavemechanics formulation in position space was not the first one to be developed. Instead, the *matrix mechanics* formulation of QM was the originally developed representation by Heisenberg in 1925, six months before Schrödinger developed the wavemechanics. In the matrix mechanics case, the state vector is projected down on an arbitrary, discrete, orthonormal set of basis vectors $|k\rangle$, $k = \{1, 2, \ldots\}$. A vector $|a\rangle$ may then be expanded as $|a\rangle = \sum_k a_k|k\rangle$ where $a_k = \langle k|a\rangle$. These coefficients can be visualized as components of a vector:

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \ldots \end{bmatrix} \tag{1.51}$$

The scalar product $\langle b|a\rangle$ is then

$$\langle b|a\rangle = \sum_k \langle b|k\rangle\langle k|a\rangle = \sum_k \langle k|b\rangle^*\langle k|a\rangle = \sum_k b_k^* a_k. \tag{1.52}$$

We see that

$$\boldsymbol{b}^\dagger \boldsymbol{a} = [b_1^*,\ b_2^*,\ \ldots] \begin{bmatrix} a_1 \\ a_2 \\ \ldots \end{bmatrix} = \sum_k b_k^* a_k. \tag{1.53}$$

With this representation, an operator $\hat{A}$ has an expectation value which is a matrix with elements $A_{mn} = \langle m|\hat{A}|n\rangle$. If $|b\rangle = \hat{A}|a\rangle$, we then obtain that

$$\langle m|b\rangle = \langle m|\hat{A}|a\rangle = \sum_n \langle m|\hat{A}|n\rangle\langle n|a\rangle,\ m = 1, 2, \ldots \tag{1.54}$$

which in turn can be written as $b_m = \sum_n A_{mn}a_n$. But this is nothing but the very definition of matrix multiplication:

$$\begin{bmatrix} b_1 \\ b_2 \\ \ldots \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \ldots \\ A_{21} & A_{22} & \ldots \\ \ldots & \ldots & \ldots \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \ldots \end{bmatrix} \tag{1.55}$$

In effect, *the result of acting with the operator on the state vector is represented by conventional matrix multiplication.* This representation is commonly used and its most important application is on the stationary SE $\hat{H}|\psi\rangle = E|\psi\rangle$. If we know the eigenvalues $|n\rangle$ (although usually we do not: the task is to find them), using them as basis vectors gives:

$$\langle m|\hat{H}|n\rangle = E_n\langle m|n\rangle \to H_{mn} = E_n\delta_{mn}. \tag{1.56}$$

We used that $\hat{H}|n\rangle = E_n|n\rangle$. The matrix-representation of $\hat{H}$ is then diagonal. Explicitly, we have

$$\begin{bmatrix} E_1 & 0 & \ldots \\ 0 & E_2 & \ldots \\ \ldots & \ldots & \ldots \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \ldots \end{bmatrix} = E \begin{bmatrix} a_1 \\ a_2 \\ \ldots \end{bmatrix}. \tag{1.57}$$

The solution for the eigenstates becomes $a_n = \delta_{mn}$, $E = E_n$. However, if the eigenvectors are *not known*, one has to use a different basis set for which $\langle m|\hat{H}|n\rangle$ is not diagonal to begin with, i.e.:

$$\begin{bmatrix} H_{11} & H_{12} & \ldots \\ H_{21} & H_{22} & \ldots \\ \ldots & \ldots & \ldots \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \ldots \end{bmatrix} = E \begin{bmatrix} c_1 \\ c_2 \\ \ldots \end{bmatrix}. \tag{1.58}$$

The task to solve the SE is then mathematically equivalent to changing the basis, $c_k = \sum_k S_{nk}a_k$, so that the matrix becomes diagonal. This is a standard method, suitable for numerics, which we later will use for degenerate perturbation theory.

E. Briefly about the Schrödinger- and Heisenberg-picture

So far, we have described quantum mechanical systems by a state $|\psi\rangle$ which "moves" in a Hilbert space where the axes (basis vectors) are time-independent. This is known as the Schrödinger picture. However, it is fully possible to take the perspective from a rotating coordinate system. The simplest option is in fact that the rotation of the system is such that the state vector is at rest. This is known as the Heisenberg picture.

Let us first recap how time-evolution is treated in the position representation. Since the SE is linear and 1st order in time, the propagator $U = U(\boldsymbol{r}, t; \boldsymbol{r}_0, t_0)$ determines the evolution of the wavefunction from t_0 to t:

$$\Psi(\boldsymbol{r}, t) = \int U(\boldsymbol{r}, t; \boldsymbol{r}_0, t_0)\Psi(\boldsymbol{r}_0, t_0)d\boldsymbol{r}_0. \tag{1.59}$$

For a Hamilton operator that does not depend explicitly on time, we can expand

$$\Psi(\boldsymbol{r},t) = \sum_n c_n \psi_n(\boldsymbol{r}) \mathrm{e}^{-\mathrm{i}E_n t/\hbar} \tag{1.60}$$

where c_n are determined by

$$c_n = \mathrm{e}^{\mathrm{i}E_n t_0/\hbar} \int \Psi(\boldsymbol{r}_0,t_0)\psi_n^*(\boldsymbol{r}_0) d\boldsymbol{r}_0. \tag{1.61}$$

Inserting this c_n into the expression for Ψ, we obtain an equation for the propagator:

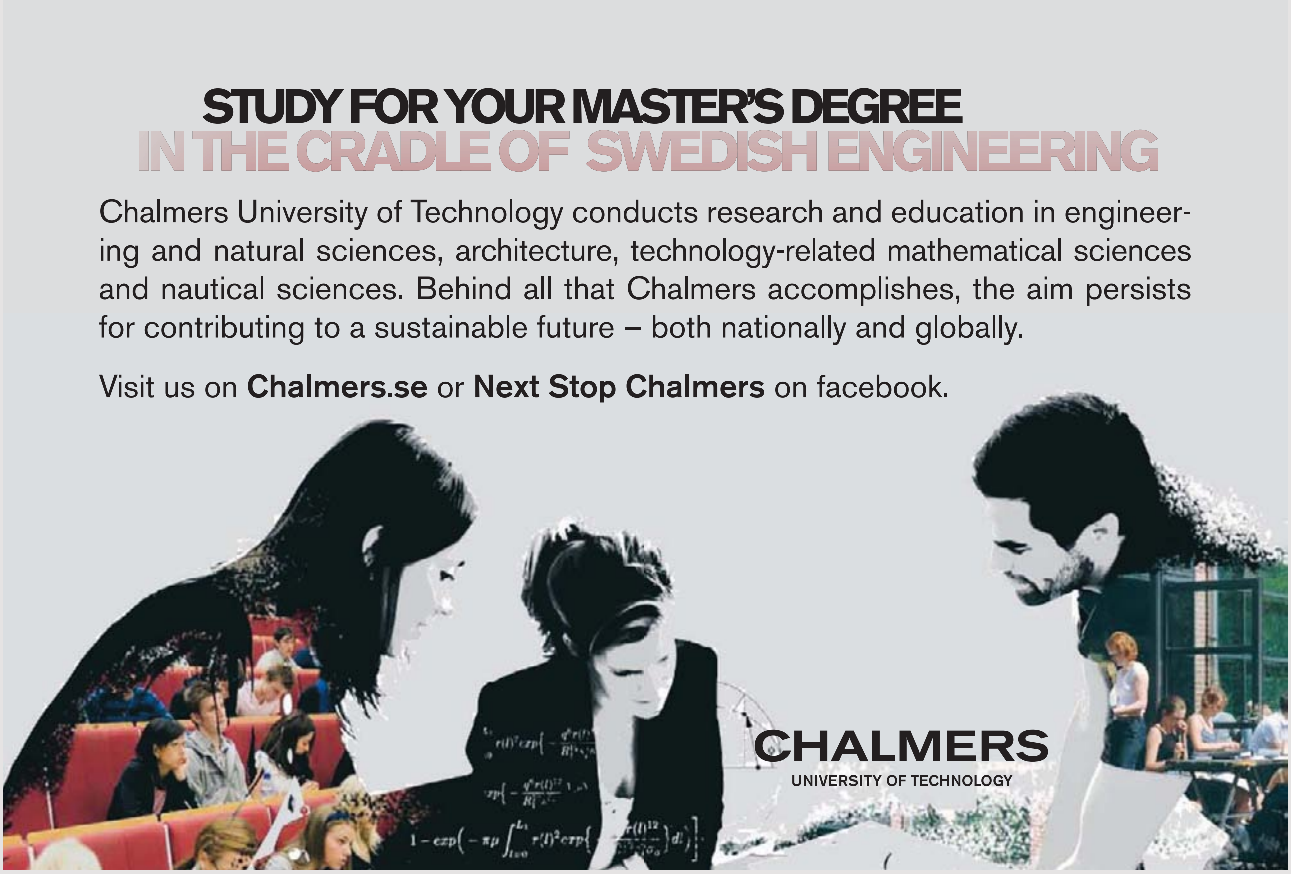
$$U(\boldsymbol{r},t;\boldsymbol{r}_0,t_0) = \sum_n \psi_n^*(\boldsymbol{r}_0)\psi_n(\boldsymbol{r}) \mathrm{e}^{-\mathrm{i}(t-t_0)E_n/\hbar}. \tag{1.62}$$

If we instead have a continuous eigenvalue spectrum, the summation is replaced by an integral:

$$U(\boldsymbol{r},t;\boldsymbol{r}_0,t_0) = \int_{-\infty}^{\infty} \psi_p^*(\boldsymbol{r}_0)\psi_p(\boldsymbol{r}) \mathrm{e}^{-\mathrm{i}(t-t_0)E_p} dp \tag{1.63}$$

where p is the eigenvalue parameter. To be concrete, consider the example of a free one-dimensionally moving particle for which

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} \mathrm{e}^{\mathrm{i}px/\hbar}, \; E_p = p^2/2m. \tag{1.64}$$



The propagator then turns into:

$$\begin{aligned} U(x,t;x_0,t_0) &= \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}p(x-x_0)/\hbar}\mathrm{e}^{-\mathrm{i}(t-t_0)p^2/2m\hbar}dp \\ &= \sqrt{\frac{m}{2\pi\mathrm{i}\hbar(t-t_0)}}\mathrm{e}^{\mathrm{i}m(x-x_0)^2/2\hbar(t-t_0)}. \end{aligned} \tag{1.65}$$

With this in mind, let us now turn to the general formulation of QM. The time evolution is given by

$$|\Psi(t)\rangle = \hat{U}(t,t_0)|\Psi(t_0)\rangle. \tag{1.66}$$

If the Hamilton operator does not contain time explicitly, we have

$$\hat{U}(t,t_0) = \mathrm{e}^{-\mathrm{i}(t-t_0)\hat{H}/\hbar} \tag{1.67}$$

where the exponential operator should be interpreted via the formula

$$\mathrm{e}^{\hat{X}} = \sum_{n=0}^{\infty}\frac{\hat{X}^n}{n!}. \tag{1.68}$$

We see that $|\Psi(t)\rangle = \mathrm{e}^{-\mathrm{i}(t-t_0)\hat{H}/\hbar}|\Psi(t_0)\rangle$ satisfies the time-dependent SE $\mathrm{i}\partial_t|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle$. The corresponding bra to the above ket is

$$\langle\Psi(t)| = \langle\Psi(t_0)|\mathrm{e}^{\mathrm{i}(t-t_0)\hat{H}/\hbar}. \tag{1.69}$$

Normalization is thus preserved since:

$$\langle\Psi(t)|\Psi(t)\rangle = \langle\Psi(t_0)|\Psi(t_0)\rangle. \tag{1.70}$$

We may compute the expectation value of some physical quantity F at the time t in the usual way:

$$\langle F\rangle = \langle\Psi(t)|\hat{F}|\Psi(t)\rangle = \langle\Psi(t_0)|\hat{F}_H|\Psi(t_0)\rangle \tag{1.71}$$

where we defined

$$\hat{F}_H = \mathrm{e}^{\mathrm{i}(t-t_0)\hat{H}/\hbar}F\mathrm{e}^{-\mathrm{i}(t-t_0)\hat{H}/\hbar}. \tag{1.72}$$

We see that $\langle F\rangle$ can be expressed in two equivalent ways:

- **Schrödinger picture**: expectation value of a time-independent operator $\hat{F}$ in a time-dependent state.
- **Heisenberg picture**: expectation value of a time-dependent operator $\hat{F}_H$ in a time-independent state.

We see that $\hat{F}_H = \hat{U}^\dagger\hat{F}\hat{U}$ where the evolution operator satisfies $\hat{U}^\dagger = \hat{U}^{-1}$, meaning that it is a unitary operator. In the Schrödinger picture, we know that

$$\frac{d}{dt}\langle F\rangle = \frac{\mathrm{i}}{\hbar}\langle[\hat{H},\hat{F}]\rangle. \tag{1.73}$$

In the Heisenberg picture, we may differentiate $\hat{F}_H$ to obtain the equation:

$$\frac{d}{dt}\hat{F}_H = \frac{\mathrm{i}}{\hbar}[\hat{H},\hat{F}_H]. \tag{1.74}$$

Note that the commutator relations are preserved when making a transition to time-dependent operators. If $[\hat{A},\hat{B}] = \hat{C}$, then

$$[\hat{A}_H,\hat{B}_H] = [\hat{U}^\dagger\hat{A}\hat{U},\hat{U}^\dagger\hat{B}\hat{U}] = \hat{U}^\dagger(\hat{A}\hat{B}-\hat{B}\hat{A})\hat{U} = \hat{U}^\dagger\hat{C}\hat{U} = \hat{C}_H. \tag{1.75}$$

Example 1. Heisenberg picture representation of creation and annihilation operators. For the creation and annihilation operators $a^\dagger$ and a of a harmonic oscillator, we obtain from Eq. (1.74):

$$\frac{da_H}{dt} = -\mathrm{i}\omega a_H,\ \frac{da_H^\dagger}{dt} = \mathrm{i}\omega a_H^\dagger. \tag{1.76}$$

The solution is straightforward to obtain:

$$a_H(t) = \mathrm{e}^{-\mathrm{i}\omega t}a_H(0),\ a_H^\dagger(t) = \mathrm{e}^{\mathrm{i}\omega t}a_H^\dagger(0). \tag{1.77}$$

It is also worth mentioning that it is possible with an approach where only part of the time dependence is transferred to the operators. This is the *interaction picture*, which is often used when the Hamilton operator can be written as $\hat{H} = \hat{H}_0 + \hat{H}_I$, where $\hat{H}_I$ has to be handled via perturbation theory. We may then transform with $\hat{U}_0 = \mathrm{e}^{-\mathrm{i}(t-t_0)\hat{H}_0/\hbar}$ so that the state vector would be time independent if $\hat{H}_I$ could be neglected. Which picture that one ultimately decides to use is a matter of convenience: the physics is the same.

II. HARMONIC OSCILLATOR: CREATION AND ANNIHILATION OPERATORS & COHERENT STATES

Learning goals. After reading this chapter, the student should:

- Be able to work with creation and annihilation operators and know their mathematical properties.
- Know how to describe coherent states and why they are physically significant.

In introductory QM courses, one learns about a wavemechanical treatment of the 1D harmonic oscillator. Let us now use operator-algebra for the state vectors in Hilbert space to study the same problem in a simpler and more elegant manner.

A. Creation and annihilation operators

The Hamilton operator for a harmonic oscillator is known from introductory courses on QM, namely:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m}{2}\omega^2\hat{q}^2. \tag{2.1}$$

It follows that the equation

$$\frac{\hat{H}}{\hbar\omega} = \frac{\hat{p}^2}{2m\hbar\omega} + \frac{m\omega}{2\hbar}\hat{q}^2 \tag{2.2}$$

is dimensionless since $[\hbar\omega]$ = energy. Instead of $\hat{q}$ and $\hat{p}$, we now introduce the dimensionless operators a and $a^\dagger$ (dropping the superscript $.\hat{.}.$ for brevity of notation)

$$a = \sqrt{\frac{m\omega}{2\hbar}}\hat{q} + \frac{\mathrm{i}}{\sqrt{2m\hbar\omega}}\hat{p},\ a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\hat{q} - \frac{\mathrm{i}}{\sqrt{2m\hbar\omega}}\hat{p}. \tag{2.3}$$

While $\hat{q}$ and $\hat{p}$ are Hermitian operators, we see that a and $a^\dagger$ are not since $a \neq a^\dagger$. It is also useful to note the inverse relations

$$\hat{q} = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger),\ \hat{p} = \mathrm{i}\sqrt{\frac{m\hbar\omega}{2}}(a^\dagger - a). \tag{2.4}$$

Keep in mind that $[\hat{q}, \hat{p}] = \mathrm{i}\hbar$. It then follows from Eq. (2.4) that

$$a^\dagger a = \frac{m\omega}{2\hbar}\hat{q}^2 + \frac{1}{2m\hbar\omega}\hat{p}^2 + \frac{\mathrm{i}}{2\hbar}(\hat{q}\hat{p} - \hat{p}\hat{q}) = \frac{\hat{H}}{\hbar\omega} - \frac{1}{2}. \tag{2.5}$$

Similarly, one shows that $aa^\dagger = \hat{H}/\hbar\omega + 1/2$. Combining these results, one obtains

$$[a, a^\dagger] = aa^\dagger - a^\dagger a = 1. \tag{2.6}$$

We have now found a very simple expression for $\hat{H}$:

$$\hat{H} = \hbar\omega(a^\dagger a + \frac{1}{2}). \tag{2.7}$$

The next step is to find the eigenvalues of $\hat{H}$. This amounts to finding the eigenvalues of $\hat{N} \equiv a^\dagger a$, since $\hat{H} = (\hat{N} + \frac{1}{2})\hbar\omega$. The quantity $\hat{N}$ is known as the *number operator*, the reason being that the eigenvalues of $\hat{N}$ are positive integers. We will now prove this. The following relations will be useful in order to accomplish this task:

$$\begin{aligned} [\hat{N}, a] &= a^\dagger aa - aa^\dagger a = (a^\dagger a - aa^\dagger)a = -a, \\ [\hat{N}, a^\dagger] &= a^\dagger aa^\dagger - a^\dagger a^\dagger a = a^\dagger(aa^\dagger - a^\dagger a) = a^\dagger. \end{aligned} \tag{2.8}$$

For reasons that will become clear soon, a is known as the annihilation operator while $a^\dagger$ is the creation operator. To identitfy the energy spectrum, let $|n\rangle$ be the orthonormal eigenvectors for $\hat{H}$ with eigenvalues E_n, so that $\hat{H}|n\rangle = E_n|n\rangle$. To find E_n, let us start by examining $a|n\rangle$. Using the above relations, we find that

$$\hat{H}a|n\rangle = a\hat{H}|n\rangle - \hbar\omega a|n\rangle = (E_n - \hbar\omega)a|n\rangle. \tag{2.9}$$

We thus conclude that as long as $a|n\rangle \neq 0$, $a|n\rangle$ is an eigenvector of $\hat{H}$ with eigenvalue $E_n - \hbar\omega$. This argument can be generalized: $a^2|n\rangle$ has eigenvalue $E_n - 2\hbar\omega$ and so forth. This cannot continue forever, however, since the energy of a harmonic oscillator cannot be negative. To see this, recall that the norm of a vector is always ≥ 0, and we have that:

$$||a|n\rangle||^2 = \langle n|a^\dagger a|n\rangle = \langle n|\frac{\hat{H}}{\hbar\omega} - \frac{1}{2}|n\rangle = \frac{E_n}{\hbar\omega} - \frac{1}{2}. \tag{2.10}$$

Therefore, we must have $E_n \geq \hbar\omega/2$. In order to guarantee that this is the case, there must exist a final eigenvector $|0\rangle$ so that $a|0\rangle = 0$. The belonging energy to the state $|0\rangle$ must be the lowest energy available, so that

$$\hat{H}|0\rangle = \hbar\omega(a^\dagger a + \frac{1}{2})|0\rangle = \frac{1}{2}\hbar\omega|0\rangle \to E_0 = \frac{1}{2}\hbar\omega. \tag{2.11}$$

Now, since we could reach this state from any higher-energy state by moving downwards with energy steps of $\hbar\omega$, we conclude that the general eigenvalues must be

$$E_n = (n + \frac{1}{2})\hbar\omega. \tag{2.12}$$

This is consistent with the known result derived in a more complicated way in introductory courses of QM, but we managed to find it in a quite simple and elegant manner using the general formulation of QM.

Let us then turn to the eigenvectors. First, note that since $\hat{H} = (\hat{N} + \frac{1}{2})\hbar\omega$ and $E_n = (n + \frac{1}{2})\hbar\omega$, it follows that $\hat{N}|n\rangle = n|n\rangle$. The eigenvalue of $\hat{N}$ thus denotes by how many energy quanta $\hbar\omega$ that the energy of the system exceeds the ground-state (lowest energy). We have that $\hat{H}a|n\rangle = (n - \frac{1}{2})\hbar\omega a|n\rangle$. But since $(n - \frac{1}{2})\hbar\omega$ is the eigenvalue of the state $|n-1\rangle$, we must have $|n-1\rangle = c_n a|n\rangle$. Here, c_n is a constant which we can determine through normalization:

$$1 = \langle n-1|n-1\rangle = |c_n|^2\langle n|a^\dagger a|n\rangle = |c_n|^2\langle n|\hat{N}|n\rangle = |c_n|^2 n. \tag{2.13}$$

Therefore, $c_n = \mathrm{e}^{\mathrm{i}\delta}/\sqrt{n}$ where $\delta \in \Re$. We set $\delta = 0$ for now and thus obtain the central result

$$a|n\rangle = \sqrt{n}|n-1\rangle. \tag{2.14}$$

However, if time-dependence is included in the notation $|n\rangle$ for stationary states, then

$$|n\rangle \propto \mathrm{e}^{-\mathrm{i}E_n t/\hbar} = \mathrm{e}^{-\mathrm{i}(n+1/2)\omega t} \tag{2.15}$$

which means that δ becomes time-dependent:

$$a|n\rangle = \mathrm{e}^{-\mathrm{i}\omega t}\sqrt{n}|n-1\rangle. \tag{2.16}$$

To find $a^\dagger|n\rangle$, we operate on the above equation on both sides with $a^\dagger$ to find

$$\mathrm{e}^{-\mathrm{i}\omega t}\sqrt{n}a^\dagger|n-1\rangle = a^\dagger a|n\rangle = \hat{N}|n\rangle = n|n\rangle, \tag{2.17}$$

which after rearranging the equation produces

$$a^\dagger|n\rangle = \mathrm{e}^{\mathrm{i}\omega t}\sqrt{n+1}|n+1\rangle. \tag{2.18}$$

This time-dependence is disregarded in the rest of this section, which means we set $t = 0$. Summarizing so far, we have then found the following two fundamental relations regarding how annihilation and creation operators act:

$$\boxed{a|n\rangle = \sqrt{n}|n-1\rangle,\ a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle.}$$

We can finally understand why $a^\dagger$ is referred to as a creation operator, since its effect is $|n\rangle \to |n+1\rangle$ (creates one quantum of energy). In the same way, a is the annihilation operator since $|n\rangle \to |n-1\rangle$. Any excited state $|n\rangle$ can thus be obtained by acting on the ground state $|0\rangle$ n times with $a^\dagger$:

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle. \tag{2.19}$$

Example 2. Computation of the expectation value of the potential energy of a harmonic oscillator in the state $|n\rangle$. We know that

$$V = \frac{1}{2}m\omega^2 q^2 = \frac{1}{2}m\omega^2 \frac{\hbar}{2m\omega}(a + a^\dagger)^2. \tag{2.20}$$

Inserted into $\langle n|V|n\rangle$, we obtain

$$\begin{aligned}\langle n|\frac{1}{4}\hbar\omega(aa + aa^\dagger + a^\dagger a + a^\dagger a^\dagger)|n\rangle &= \langle n|\frac{1}{4}\hbar\omega(aa^\dagger + a^\dagger a)|n\rangle \\ &= \frac{1}{4}\hbar\omega(\sqrt{n+1}\sqrt{n+1} + \sqrt{n}\sqrt{n}) \\ &= \frac{1}{2}\hbar\omega(n + 1/2) = E_n/2. \end{aligned} \tag{2.21}$$

The average potential energy is thus equal to the average kinetic energy, namely 50% of the total energy in state $|n\rangle$. Note that, in comparison, if we wanted to compute $\langle q^2\rangle$ in the position representation, it would have been necessary to evluate an integral with the square of a Hermite-polynomial, which is a much more difficult task!

For completeness, let us show how the position representation wavefunctions are recovered from the eigenstates $|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle$. We know that the starting point to find $\psi_n(q)$, where q is the position coordinate, is $\psi_n(q) = \langle q|n\rangle$. Using the completeness relation $\int dq'|q'\rangle\langle q'| = 1$, we obtain

$$\langle q|n\rangle = \frac{1}{\sqrt{n!}}\int dq'\langle q|(a^\dagger)^n|q'\rangle\langle q'|0\rangle. \tag{2.22}$$

First, we evaluate

$$\begin{aligned}\langle q|(a^\dagger)^n|q'\rangle &= \langle q|\Big(\sqrt{\frac{m\omega}{2\hbar}}\hat{q} - \frac{\mathrm{i}}{\sqrt{2m\hbar\omega}}\hat{p}\Big)^n|q'\rangle \\ &= \Big(\sqrt{\frac{m\omega}{2\hbar}}\hat{q} - \frac{\hbar}{\sqrt{2m\hbar\omega}}\frac{d}{dq}\Big)^n\delta(q-q'). \end{aligned} \tag{2.23}$$

Inserting this into Eq. (2.22), we obtain

$$\langle q|n\rangle = \frac{1}{\sqrt{n!}}\Big(\sqrt{\frac{m\omega}{2\hbar}}\hat{q} - \frac{\hbar}{\sqrt{2m\hbar\omega}}\frac{d}{dq}\Big)^n\langle q|0\rangle. \tag{2.24}$$

We see that the n-th wavefunction $\psi_n = \langle q|n\rangle$ is expressed via $\psi_0(q) = \langle q|0\rangle$. We determine $\psi_0(q)$ by the criterion that defined $|0\rangle$, namely $a|0\rangle = 0$. Projected onto $|q\rangle$, we get:

$$\langle q|a|0\rangle = \int dq'\langle q|a|q'\rangle\langle q'|0\rangle = 0. \tag{2.25}$$

In turn, this yields

$$\begin{aligned}\langle q|a|0\rangle = 0 &= \int dq'\langle q|\sqrt{\frac{m\omega}{2\hbar}}\hat{q} + \frac{\mathrm{i}}{\sqrt{2m\hbar\omega}}\hat{p}|q'\rangle\langle q'|0\rangle \\ &= \Big(\sqrt{\frac{m\omega}{2\hbar}}\hat{q} + \frac{\hbar}{\sqrt{2m\hbar\omega}}\frac{d}{dq}\Big)\langle q|0\rangle. \end{aligned} \tag{2.26}$$

This means that we have obtained the following differential equation for the scalar $\langle q|0\rangle$:

$$\frac{d}{dq}\langle q|0\rangle = -\frac{m\omega}{\hbar}q\langle q|0\rangle. \tag{2.27}$$

It can be readily solved to yield $\ln\langle q|0\rangle = -\frac{m\omega}{2\hbar}q^2 + C$ where C is a constant. Therefore,

$$\langle q|0\rangle = \mathrm{e}^C\mathrm{e}^{-m\omega q^2/2\hbar} = \Big(\frac{m\omega}{\pi\hbar}\Big)^{1/4}\mathrm{e}^{-m\omega q^2/2\hbar} \tag{2.28}$$

where we determined the constant through normalization. Now, by inserting this back into our expression for $\psi_n(q)$ we get:

$$\psi_n = \Big(\frac{m\omega}{\pi\hbar}\Big)^{1/4}\frac{1}{\sqrt{2^n n!}}\Big(x - \frac{d}{dx}\Big)^n \mathrm{e}^{-x^2/2} \tag{2.29}$$

where $x \equiv q\sqrt{m\omega/\hbar}$, which is the correct result for the position representation wavefunction.

We mention in passing that we can now also identify the matrix representation for the operators by using the energy eigenvectors as a basis. From $a|n\rangle = \sqrt{n}|n-1\rangle$ and $a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$, we see that

$$a = \begin{bmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{2} & 0 & \dots \\ 0 & 0 & 0 & \sqrt{3} & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix},\ a^\dagger = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ \sqrt{1} & 0 & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \tag{2.30}$$

B. Coherent states

The eigenstates of the annihilation operator a, $|\alpha\rangle$, are known as coherent states:

$$a|\alpha\rangle \propto |\alpha\rangle. \tag{2.31}$$

The reason for this is that the time-evolution of such a state does not cause the state to spatially "diffuse" and become delocalized. Instead, the state's spatial distribution oscillates with a preserved width of the oscillation as we now shall prove.

We expand the eigenstate of the operator a in energy eigenstates:

$$|\alpha\rangle = \sum_{n=0}^{\infty} c_n|n\rangle. \tag{2.32}$$

We showed previously that:

$$a|\alpha\rangle = \mathrm{e}^{-\mathrm{i}\omega t}\sum_{n=0}^{\infty} c_n\sqrt{n}|n-1\rangle. \tag{2.33}$$

When $c_n\sqrt{n} = \alpha c_{n-1}$, where α is a constant, $a|\alpha\rangle$ becomes proportional to $|\alpha\rangle$. Using this relation, we have $c_n = c_0\alpha^n/\sqrt{n!}$, so that

$$|\alpha\rangle = c_0\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle = \mathrm{e}^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle \tag{2.34}$$

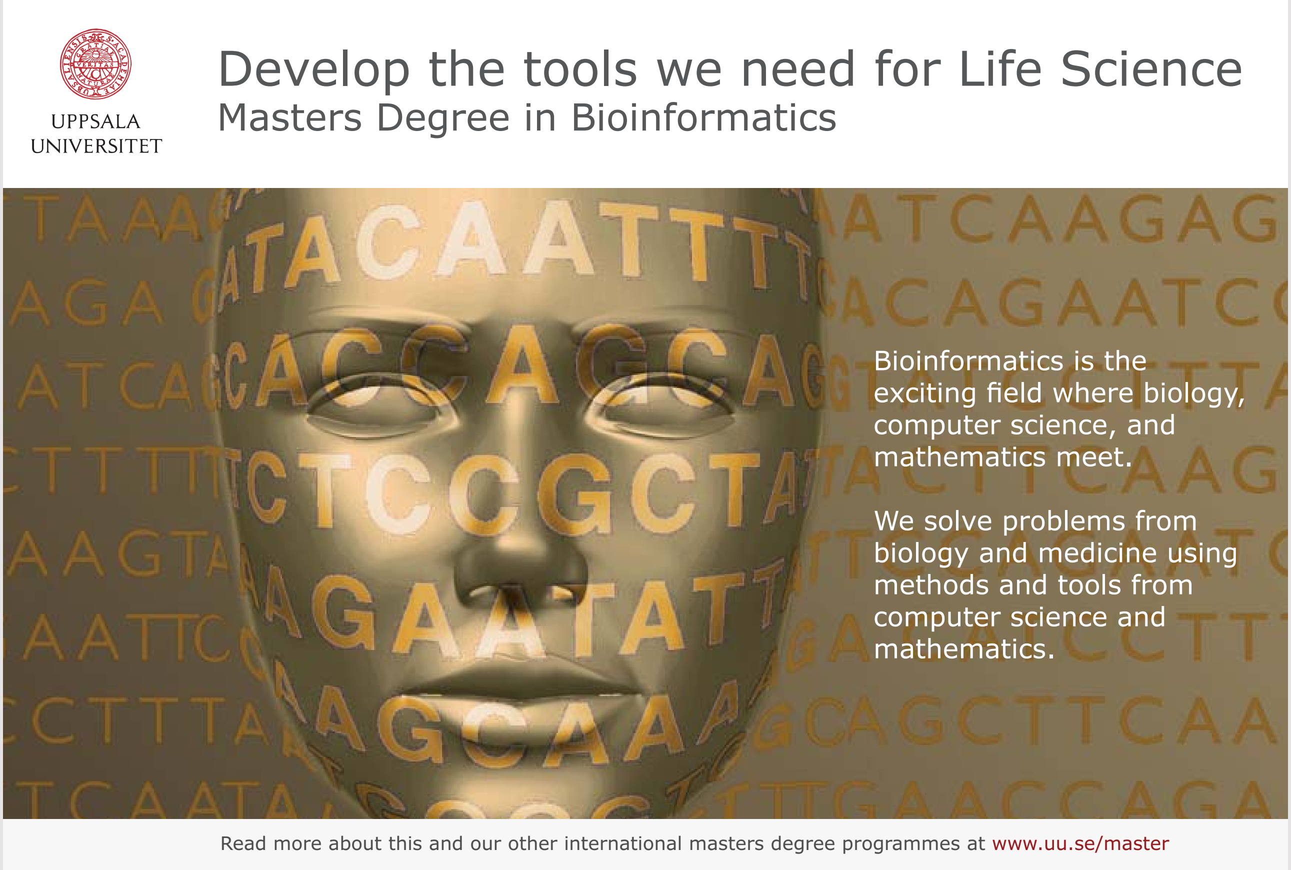
We have chosen c_0 so that $\langle\alpha|\alpha\rangle = 1$ by using that $\sum_{n=0}^{\infty}(\alpha^*\alpha)^n/n! = \mathrm{e}^{|\alpha|^2}$. These states then satisfy $a|\alpha\rangle = \mathrm{e}^{-\mathrm{i}\omega t}\alpha|\alpha\rangle$ and the expectation values for a and $a^\dagger$ are:

$$\langle\alpha|a|\alpha\rangle = \alpha\mathrm{e}^{-\mathrm{i}\omega t},\ \langle\alpha|a^\dagger|\alpha\rangle^* = \alpha^*\mathrm{e}^{\mathrm{i}\omega t}. \tag{2.35}$$

It remains to justify why we have said that these are known as *coherent states*. In order to see this, we consider how these states behave spatially. Since $\hat{q} = \sqrt{\hbar/2m\omega}(a + a^\dagger)$, we can show that:

$$|\langle q|\alpha\rangle|^2 = \sqrt{\frac{m\omega}{\pi\hbar}}\mathrm{e}^{-m\omega[q-q_0\cos(\omega t-\theta)]^2/\hbar}, \tag{2.36}$$

where $\alpha = |\alpha|\mathrm{e}^{\mathrm{i}\theta}$. The meaning of this inner product is the distribution of the spatial position, which is seen to describe an oscillating wavepacket which maintains a constant width as time evolves, hence the name coherent state.



III. TIME-INDEPENDENT APPROXIMATE METHODS

Learning goals. After reading this chapter, the student should:

- Know the fundamental idea behind non-degenerate and degenerate perturbation theory, when they are valid, and be able to mathematically outline how to apply them on a quantum mechanical problem..
- Know the fundamental idea behind the variational method, when it is valid, and be able to mathematically outline how to apply it on a quantum mechanical problem.
- Know the fundamental idea behind the WKB-approximation, when it is valid, and be able to mathematically outline how to apply it on a quantum mechanical problem.

Only rarely is a QM problem exactly solvable. Thus, having a "toolbox" of useful approximative methods is indispensible for a physicist. In this chapter, we will establish precisely such a toolbox.

A. Non-degenerate perturbation theory

Assume that $\hat{H}_0$ corresponds to an exactly solvable problem. Often times, a physical system may be described by a $\hat{H}$ which only slightly deviates from $\hat{H}_0$. Then, $\hat{H}-\hat{H}_0$ is the perturbation of the system. Assume E_n^0 and $|n\rangle$ are known for $\hat{H}_0$

$$\hat{H}_0|n\rangle = E_n^0|n\rangle. \tag{3.1}$$

We want to find eigenvalues and eigenstates for the perturbed Hamilton operator $\hat{H}=\hat{H}_0+\lambda\hat{H}_1$, where $\hat{H}_1$ is time independent. Here, λ is an expansion parameter which is assumed to be small. This kind of perturbation theory is suitable and commonly used in the context of atomic energy levels influenced by $\boldsymbol{E}$ or $\boldsymbol{B}$ fields. We start by assuming that the unperturbed energy level is non-degenerate. The exact eigenvalue problem can then be written as:

$$(\hat{H}_0+\lambda\hat{H}_1-E_n)|\psi_n\rangle = 0. \tag{3.2}$$

We now expand the eigenvalues and eigenstates in corrections to the unperturbed solutions:

$$\begin{aligned} E_n &= E_n^0+\lambda E_n^{(1)}+\lambda^2E_n^{(2)}+\ldots, \\ |\psi_n\rangle &= |n^{(0)}\rangle+\lambda|n^{(1)}\rangle+\ldots \end{aligned} \tag{3.3}$$

and thus obtain

$$(\hat{H}_0+\lambda\hat{H}_1-E_n^0-\lambda E_n^{(1)}-\ldots)(|n^{(0)}\rangle+\lambda|n^{(1)}\rangle+\ldots)=0. \tag{3.4}$$

For brevity of notation, we use $|n\rangle\equiv|n^{(0)}\rangle$ in what follows. If this is to be valid for all λ, the equation must be fulfilled for each power of λ. We obtain to $\mathcal{O}(\lambda^0)$:

$$(\hat{H}_0-E_n^{(0)})|n\rangle = 0, \tag{3.5}$$

while to order $\mathcal{O}(\lambda^1)$:

$$(\hat{H}_0-E_n^{(0)})|n^{(1)}\rangle+(\hat{H}_1-E_n^{(1)})|n\rangle = 0 \tag{3.6}$$

and finally to order $\mathcal{O}(\lambda^2)$:

$$(\hat{H}_0-E_n^{(0)})|n^{(2)}\rangle+(\hat{H}_1-E_n^{(1)})|n^{(1)}\rangle-E_n^{(2)}|n\rangle = 0. \tag{3.7}$$

The 0th order equation is known to be valid from the outset, since it corresponds to the exact unperturbed problem. If we multiply the 1st order equation $\langle n|$ from the left we obtain:

$$\langle n|\hat{H}_0-E_n^0|n^{(1)}\rangle+\langle n|\hat{H}_1|n\rangle = E_n^{(1)}. \tag{3.8}$$

The first term is zero since it is equal to $\langle n^{(1)}|\hat{H}_0-E_n^0|n\rangle^* = 0$. Therefore, we obtain

$$\boxed{\lambda E_n^{(1)} = \langle n|\lambda \hat{H}_1|n\rangle.}$$

This is the lowest order energy correction. It can also be written explicitly as

$$\lambda E_n^{(1)} = \int [\psi_n^{(0)}]^* \lambda \hat{H}_1 \psi_n^{(0)} d\boldsymbol{r} \tag{3.9}$$

What about the correction to the eigenstates? Multiply the $\mathcal{O}(\lambda)$ equation with $\langle m|$ where $m \neq n$ to obtain

$$\langle m|\hat{H}_0 - E_n^{(0)}|n^{(1)}\rangle + \langle m|\hat{H}_1|n\rangle = 0. \tag{3.10}$$

Defining $E_j^0 \equiv E_j^{(0)}$ and using that

$$\begin{aligned}\langle m|\hat{H}_0 - E_n^0|n^{(1)}\rangle &= \langle n^{(1)}|\hat{H}_0 - E_n^0|m\rangle^* \\ &= (E_m^0 - E_n^0)\langle n^{(1)}|m\rangle^* \\ &= (E_m^0 - E_n^0)\langle m|n^{(1)}\rangle,\end{aligned} \tag{3.11}$$

we obtain

$$\langle m|n^{(1)}\rangle = \frac{\langle m|\hat{H}_1|n\rangle}{E_n^0 - E_m^0}. \tag{3.12}$$

It is now clear why a problem would arise if the unperturbed eigenvalues were degenerate, since the denominator would beome zero then. By finally expanding $|n^{(1)}\rangle$ in the unperturbed eigenstates

$$|n^{(1)}\rangle = \sum_m |m\rangle\langle m|n^{(1)}\rangle \tag{3.13}$$

via the completeness relation, we end up with the first order correction to the eigenstates:

$$|n^{(1)}\rangle = \sum_{m\neq n} \frac{\langle m|\hat{H}_1|n\rangle}{E_n^0 - E_m^0}|m\rangle. \tag{3.14}$$

We have now determined the eigenvalues and eigenstates up to $\mathcal{O}(\lambda^1)$. For some applications, it turns out that $\langle n|\hat{H}_1|n\rangle = 0$, which means that we have to go to second order in λ to find the first non-vanishing correction. Following a similar procedure as in the first order case, one obtains for the eigenvalues

$$E_n = E_n^0 + \langle n|\lambda\hat{H}_1|n\rangle + \sum_{m\neq n} \frac{|\langle m|\lambda\hat{H}_1|n\rangle|^2}{E_n^0 - E_m^0} + \ldots \tag{3.15}$$

Note that if we are perturbing the ground state, then $E_m^0 > E_n^0$, which means that $E_n^{(2)}$ will always be negative. Moreover, the above expression gives a criterion for the applicability of this method, namely that

$$|\langle m|\lambda\hat{H}_1|n\rangle| \ll |E_n^0 - E_m^0| \tag{3.16}$$

so that the correction to E_n is indeed small as assumed. This type of approximation theory is known as Rayleigh-Schrödinger perturbation theory.

Example 3. Relativistic correction to the Coulomb-levels. Even if the levels above the ground-state have a degeneracy, we can still use our approximation theory because the perturbation matrix elements between degenerate states, $\langle m|\lambda\hat{H}_1|n\rangle$, turn out to vanish. The relativistic expression for kinetic energy can be expandaed in momentum as follows:

$$\sqrt{m^2c^4 + \boldsymbol{p}^2c^2} - mc^2 = mc^2\sqrt{1 + \boldsymbol{p}^2/m^2c^2} - mc^2 \tag{3.17}$$

and when assuming that $|\boldsymbol{p}| \ll mc$, we obtain

$$mc^2\sqrt{1 + \boldsymbol{p}^2/m^2c^2} - mc^2 \simeq \frac{\boldsymbol{p}^2}{2m} - \frac{\boldsymbol{p}^4}{8m^3c^2} + \ldots \tag{3.18}$$

Here, m is the rest mass of the particle. The perturbation is then:

$$\lambda \hat{H}_1 = -\frac{\hbar^4}{8m^3c^2}\nabla^4. \tag{3.19}$$

We know that the first order correction to the energy eigenvalue is the expectation value:

$$\lambda E_n^{(1)} = \langle n|\lambda \hat{H}_1|n\rangle = -\frac{\hbar^4}{8m^3c^2}\int |\nabla^2\psi_n|^2 d\boldsymbol{r}. \tag{3.20}$$

This result was obtained by performing two partial integrations. The integral is most easily evaluated by using the SE for the Coulomb-potential which reads

$$-\frac{\hbar^2}{2m}\nabla^2\psi_{nlm} = \frac{Ze^2}{4\pi\epsilon_0}\psi_{nlm} + E_n\psi_{nlm} = \frac{Ze^2}{4\pi\epsilon_0}\Big(\frac{1}{r} - \frac{1}{2an^2}\Big)\psi_{nlm}. \tag{3.21}$$

Inserted into our expression Eq. (3.20), we obtain

$$\lambda E^{(1)} = -\frac{1}{2mc^2}\Big(\frac{Ze^2}{4\pi\epsilon_0}\Big)^2\langle\Big(\frac{1}{r} - \frac{1}{2an^2}\Big)^2\rangle. \tag{3.22}$$

This expectation value may be computed by using the known form of the hydrogen wavefunction. Introducing the fine-structure constant $\alpha = e^2/4\pi\epsilon_0\hbar c$, we obtain

$$E_{nl} = mc^2\Big[1 - \frac{Z^2\alpha^2}{2n^2} - \frac{Z^4\alpha^4}{n^4}\Big(\frac{n}{2l+1} - \frac{3}{8}\Big)\Big]. \tag{3.23}$$

The term $\propto Z^4$ is the lowest order relativistic correction to the energy level. Importantly, the energy level is now not only dependent on n, but also on the angular momentum quantum number l. This means that the energy spectrum has acquired a *fine-structure*. This result is correct for a spinless particle. The result is slightly modified for *e.g.* an electron that has spin 1/2.

B. Degenerate perturbation theory

Consider now the case where the eigenvalue E_n^0 for the unperturbed Hamilton operator $\hat{H}^0$ is degenerate. Let there be g orthonormal states $|n_1\rangle, |n_2\rangle, \ldots$ with eigenvalue E_n^0. An example is the four states $|200\rangle$, $|211\rangle$, $|210\rangle$, $|21-1\rangle$ corresponding to the first excited level for the Coulomb potential. We again expand in powers of λ. For the state vector:

$$|\psi_n\rangle = \sum_{r=1}^{g} a_r|n_r\rangle + \lambda|N_1\rangle + \ldots \tag{3.24}$$

with so far unknown coefficients a_r. To first order in λ, the time independent SE gives:

$$(\hat{H}_0 - E_n^0)|N_1\rangle + (\hat{H}_1 - E_n^{(1)})\sum_{r=1}^{g} a_r|n_r\rangle = 0. \tag{3.25}$$

Multiplying from the left with one of the unperturbed states $\langle n_s|$:

$$\sum_{r=1}^{g}\langle n_s|\hat{H}_1 - E_n^{(1)}|n_r\rangle a_r = 0 \tag{3.26}$$

where we used that $\langle N_1|\hat{H}_0 - E_n^0|n_s\rangle^* = 0$. In Eq. (3.26), all the matrix elements are known:

$$\langle n_s|\hat{H}_1|n_r\rangle \equiv H'_{sr}. \tag{3.27}$$

They are computed via the unperturbed states. Since $\langle n_s|n_r\rangle = \delta_{sr}$, we get

$$\sum_{r=1}^{g}(H'_{sr} - E_n^{(1)}\delta_{sr})a_r = 0,\ s = 1, 2, \ldots g. \tag{3.28}$$

This is in fact a homogeneous set of equations for the unknown a_r:

$$\begin{bmatrix} H'_{11} - E_n^{(1)} & H'_{12} & \ldots & H'_{1g} \\ H'_{21} & H'_{22} - E_n^{(1)} & \ldots & H'_{2g} \\ \ldots & \ldots & \ldots & \ldots \\ H'_{g1} & H'_{g2} & \ldots & H'_{gg} - E_n^{(1)} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \ldots \\ a_g \end{bmatrix} = 0. \tag{3.29}$$

This only has a non-trivial solution for the coefficients $\{a_r\}$ when $\det(M) = 0$ where M is the matrix in the above equation. This gives an equation of the g-th degree for $E_n^{(1)}$. If all g solutions for $E_n^{(1)}$ are different, it means that the perturbation $\lambda\hat{H}_1$ has completely lifted the degeneracy of the energy level and split it into g levels. This method for degenerate levels can and *should* be used on a level which is not *exactly* degenerate, but *nearly* degenerate, so that the criterion $|\langle m|\lambda\hat{H}_1|n\rangle \ll |E_n^0 - E_m^0|$ is not safisfied. Here, $|n\rangle$ and $|m\rangle$ are unperturbed eigenstates of the Hamiltonian.

As an application of this framework, we consider the Stark-effect: the displacement of energy levels due to an external constant electric field $\boldsymbol{\mathcal{E}}$. Choosing $\boldsymbol{z}$ as the direction of the field, we get $\lambda\hat{H}_1 = e\mathcal{E}\boldsymbol{z}$ where $\boldsymbol{\mathcal{E}} = \mathcal{E}\boldsymbol{z}$. The perturbation is thus the potential energy for a charged particle in an electric field. Assume that the particle is an electron in a Coulomb potential and that the field is so weak that perturbation theory is permissible. Let the energy states in the Coulomb potential be denoted $|nlm\rangle$. The ground state $|100\rangle$ is non-degenerate, and the correction to the ground state energy E_1 becomes

$$\lambda E_1^{(1)} = e\mathcal{E}\langle 100|z|100\rangle = e\mathcal{E}\int z|\psi_{100}|^2 d\boldsymbol{r}. \tag{3.30}$$

However, this integral is zero due to symmetry since $\psi_{100} \propto \mathrm{e}^{-r/a}$. Therefore, the lowest order non-vanishing correction to the ground state is 2nd order in the perturbation (the field $\boldsymbol{\mathcal{E}}$):

$$E_1 = E_1^0 - \text{constant} \times \mathcal{E}^2. \tag{3.31}$$

The constant may be evaluated using our formula for the 2nd order correction and one finds:

$$E_1 = E_1^0\Big[1 + \frac{9}{8}\Big(\frac{ea_0\mathcal{E}}{E_1^0}\Big)^2\Big], \tag{3.32}$$

where we emphasize that $E_1^0 < 0$. The correction to the ground state is atypical, because it is quadratic in the perturbation. For all other levels, the Stark effect is actually linear in the field $\mathcal{E} = |\boldsymbol{\mathcal{E}}|$, and the reason for this is that all excited energy levels are degenerate. Consider for instance the $n = 2$ level which has a 4-fold degeneracy: $|200\rangle, |210\rangle, |211\rangle, |21-1\rangle$. To do perturbation theory for a degenerate level, we need the matrix elements $e\mathcal{E}\langle 2lm|z|2l'm'\rangle$. Several of these vanish:

- Diagonal elements $\langle 2lm|z|2lm\rangle$ are zero due to symmetry, just like the first-order term for the ground state.
- All elements where $m \neq m'$ vanish, the reason being that $z = r\cos\theta$ does not contain ϕ, while $\psi_{nlm} \propto \mathrm{e}^{\mathrm{i}m\phi}$. As a result, the ϕ-integration gives

$$\int_0^{2\pi} \mathrm{e}^{-\mathrm{i}m\phi}\mathrm{e}^{\mathrm{i}m'\phi} = 2\pi\delta_{mm'}. \tag{3.33}$$

In effect, we fortunately only need to evaluate the matrix elements $\langle 210|z|200\rangle$ and $\langle 200|z|210\rangle = \langle 210|z|200\rangle^*$. Therefore, it suffices to compute

$$e\mathcal{E}\langle 210|z|200\rangle = e\mathcal{E}\int \psi_{210}^* z\psi_{200} d\boldsymbol{r}. \tag{3.34}$$

The wavefunctions in the integral can be derived or looked up in a table and we simply write the result here:

$$\psi_{210} = (32\pi a_0^3)^{-1/2} r a_0^{-1}\mathrm{e}^{-r/2a_0}\cos\theta,\ \psi_{200} = (32\pi a_0^3)^{-1/2}(2 - r a_0^{-1})\mathrm{e}^{-r/2a_0}. \tag{3.35}$$

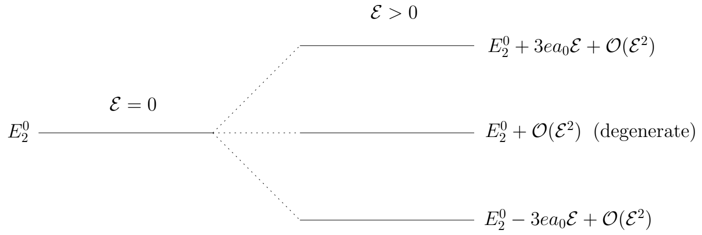
Inserted into the integral one obtains

$$e\mathcal{E}\langle 210|z|200\rangle = -3e\mathcal{E}a_0. \tag{3.36}$$

The determinant that provides us with the first order energy correction is then:

$$\begin{vmatrix} -E_2^{(1)} & -3ea_0\mathcal{E} & 0 & 0 \\ -3ea_0\mathcal{E} & -E_2^{(1)} & 0 & 0 \\ 0 & 0 & -E_2^{(1)} & 0 \\ 0 & 0 & 0 & -E_2^{(1)} \end{vmatrix} = (E_2^{(1)})^2[(E_2^{(1)})^2 - (3ea_0\mathcal{E})^2] = 0. \tag{3.37}$$

The solutions are $E_2^{(1)} = 0, 0, \pm 3ea_0\mathcal{E}$. We thus see that the degeneracy is not completely lifted: the field splits the $n = 2$ level into three levels instead of four as shown in the figure.



We here assumed that the field $\mathcal{E}$ is weak in order to use perturbation theory, but what does weak mean quantitatively? Let us compare the energy splitting due to the field with the distance to the next unperturbed energy level which is

$$E_3^0 - E_2^0 = \Big(\frac{1}{4} - \frac{1}{9}\Big)E_1^0| = 1.89 \text{ eV}. \tag{3.38}$$

The ratio between the field-splitting of the levels and the above energy gap is then

$$\frac{3ea_0\mathcal{E}}{E_3^0 - E_2^0} = \frac{\mathcal{E}}{1.2 \times 10^{10} \text{ V/m}}. \tag{3.39}$$

We may conclude that our approach is valid so long as $\mathcal{E} = |\boldsymbol{\mathcal{E}}| \ll 10^{10}$ V/m, which is an extremely large electric field. Finally, it is instructive to consider the state belonging to the lowest energy level $E_2^{(1)} = -3ea_0\mathcal{E}$. The state is specified by computing the $\{a_j\}$ coefficients in our previous derivation and one finds that

$$|\psi_-\rangle = \frac{1}{\sqrt{2}}(|200\rangle - |210\rangle). \tag{3.40}$$

This state has a finite dipole-moment along the $\boldsymbol{z}$-axis, namely

$$d = \langle\psi_-| - ez|\psi_-\rangle = 3a_0e. \tag{3.41}$$

Therefore, the physical meaning of the energy shift due to the electric field is that it represents the *dipole-energy* $-\boldsymbol{d}\cdot\boldsymbol{\mathcal{E}}$.

C. Variational method

There are problems where one cannot split $\hat{H}$ into an exactly solvable part and a small perturbation. In such events, perturbation theory is not applicable and we may instead employ a so called *variational method*. This method is particularly useful to determine the lowest-lying eigenvalue E_0. It is based on the fact that the expectation value $\hat{H}$ in any state $|f\rangle$ must be $\geq E_0$:

$$\boxed{\frac{\langle f|\hat{H}|f\rangle}{\langle f|f\rangle} \geq E_0.}$$

We prove this result using the wavemechanics formulation of QM. First, expand the states f in the eigenfunctions of $\hat{H}$, so that $f = \sum_n c_n \psi_n$. Using the orthonormality of the set $\{\psi_n\}$, we obtain $\int f^* f d\boldsymbol{r} = \sum_n |c_n|^2$. Therefore:

$$\int f^* \hat{H} f d\boldsymbol{r} = \sum_n |c_n|^2 E_n. \tag{3.42}$$

Since $E_n \geq E_0$ per definition, we obtain

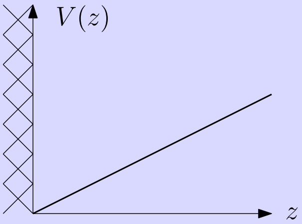
$$\int f^* \hat{H} f d\boldsymbol{r} \geq E_0 \sum_n |c_n|^2 = E_0 \int f^* f d\boldsymbol{r}, \tag{3.43}$$

which completes the proof. The equality sign is obtained only if the state f actually is the ground state ψ_0. In other words, E_0 is obtained by minimizing the functional $E[f]$ with respect to the function f where

$$E[f] = \frac{\int f^* \hat{H} f d\boldsymbol{r}}{\int f^* f d\boldsymbol{r}}. \tag{3.44}$$

The variational method then consists of selecting trial functions f that depend on one or more parameters, computing $E[f]$, and then minimizing it w.r.t. f. The result will be an upper limit for E_0, and the lowest value obtained will always be the best. To be successful, one should ideally try to guess on a trial function form f which seems physically reasonable for the system.

Example 4. Triangular well. We use the variational method to estimate E_0 for a triangular well, where $V(z) = \infty$ for $z < 0$ and $V(z) = Fz$ for $z \geq 0$.



Physically, this could be realized by pushing an electron with an electric field $\mathcal{E}$ toward a hard potential wall, so that the force is $F = e\mathcal{E}$. This is a commonly encountered situation in experimental electronics when one wants to create artificial 2D electron systems. What kind of wavefunction should we expect in this system? It should be zero upon entering the $V = \infty$ region (and in the region itself) and also fall off as z increases. Thus, a possible choice which satisfies this is $f(z) = z\mathrm{e}^{-\alpha z/2}$ for $z \geq 0$. We then obtain

$$E[f] = \frac{\int_0^\infty (-\hbar^2/2m) f f'' dz + F\int_0^\infty z f^2 dz}{\int_0^\infty f^2 dz}. \tag{3.45}$$

All integrals may be evaluated analytically, and in total one obtains

$$E[f] = \frac{\hbar^2}{2m}\frac{\alpha^2}{4} + \frac{3F}{\alpha}. \tag{3.46}$$

Here, α is the free parameter that we may adjust in order to obtain as good a guess as possible for E_0. In effect, we want to minimize $E[f]$ with respect to α. Setting $\partial E[f]/\partial\alpha = 0$ gives

$$\alpha = \left(\frac{12mF}{\hbar^2}\right)^{1/3}. \tag{3.47}$$

The corresponding minimum value of $E[f]$ for our particular trial function is then $2.48(\hbar^2/2m)^{1/3}F^{2/3}$ for that choice of α. Now, we don't know how good this result is, i.e. how far away from the true ground state energy it is. In this particular case, however, we are lucky because the triangular well problem can actually be solved exactly.

To see this, consider the SE for $z \geq 0$:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial z^2} + Fz\psi = E\psi. \tag{3.48}$$

Now, introduce the quantities $\kappa = (\hbar^2/2mF)^{1/3}$ and $x = z/\kappa$ in order to bring the equation to dimensionless form:

$$\frac{\partial^2\psi}{\partial x^2} - (x - \tilde{E})\psi = 0, \tag{3.49}$$

where we defined $\tilde{E} = E/Fk$. The key observation here is that the equation $y'' - xy = 0$ is Airy's differential equation, which has two known independent solutions: $y = \mathrm{Ai}(x)$ and $y = \mathrm{Bi}(x)$. While $\mathrm{Bi}(x)$ diverges for large x, and thus is physically unacceptable in our system, $\mathrm{Ai}(x - \tilde{E})$ has an acceptable behavior as it decreases for $x - \tilde{E} > 0$. The physically acceptable solution to the SE for this system thus has to be $\psi(x) = \mathrm{Ai}(x - \tilde{E})$ where the definition is:

$$\mathrm{Ai}(x) = \frac{1}{\pi}\int_0^\infty \cos(xz + z^3/3)dz. \tag{3.50}$$

Since $\psi(0) = 0$ due to the infinite wall potential, we obtain the energy eigenvalues from $\mathrm{Ai}(-\tilde{E}) = 0$. The smallest value of $\tilde{E}$ must be the ground state, which is found numerically to occur at $\tilde{E} = 2.33811$. Since $\tilde{E} = E/Fk$, we get

$$E_0 = 2.33811(\hbar^2/2m)^{1/3}F^{2/3}. \tag{3.51}$$

Comparing with the result we obtained using the variational method, we see now that it was quite good: only 6% deviation from the exact result!

The variational method can also be used for the lowest-lying excited level E_1, granted that we can choose a trial function that is orthogonal to the ground state. To see this, expand $f = \sum_{n=1}^{\infty} c_n\psi_n$, thus excluding $n = 0$ since f has to be orthogonal to ψ_0. It follows, proceeding as we did before, that for this f we have

$$\frac{\langle f|\hat{H}|f\rangle}{\langle f|f|\rangle} \geq E_1. \tag{3.52}$$

Symmetry can be used as a guideline to ensure orthogonality between f and ψ_0. For instance, in a 1D symmetric potential, the first excited state is antisymmetric while the ground state is symmetric. If symmetry arguments are not available, another option is to compute the ground state as accurately as we can, and then ensure that f is orthogonal to that function.

D. WKB approximation

Whereas the variatonal method is useful for approximating the ground state of a system, it is useless for the purpose of determining highly excited states. In contrast, the WKB-method (Wentzel, Kramers, Brillouin) is particularly accurate for highly excited states, and also quite accurate for lower states. This method is also known as a semiclassical approach and the key idea behind is to assume that the potential *varies slowly* in space (we will later specify what this means quantitatively, i.e. how slowly it must vary).

To outline the strategy behind the WKB approximation, consider the 1D SE with a general potential $V(x)$

$$\frac{d^2}{dx^2}\psi(x) + \frac{2m}{\hbar^2}[E - V(x)]\psi(x) = 0. \tag{3.53}$$

We now try to solve this using the *ansatz* $\psi(x) = \mathrm{e}^{\mathrm{i}S(x)/\hbar}$. For $V(x) = V_0$, this is indeed an exact solution with

$$S(x) = \pm\sqrt{2m(E - V)}x. \tag{3.54}$$

We may thus view $\psi(x)$ as a wavefunction with variable wavelength. Inserting it into the SE gives the following equation for S:

$$(S')^2 - 2m[E - V(x)] - \mathrm{i}\hbar S'' = 0. \tag{3.55}$$

If $V(x) = V_0$, then $S'' = 0$. Thus, if the potential is slowly varying, it seems reasonable to solve Eq. (3.55) iteratively while treating the term $\mathrm{i}\hbar S''$ as a small perturbation. Let us use $\hbar$ as a book-keeping expansion parameter, similarly to what we did with λ in previous perturbation theory. We expand

$$S(x) = S_0(x) + \hbar S_1(x) + \hbar^2 S_2(x) + \ldots \tag{3.56}$$

Inserting this expansion into Eq. (3.55), we first collect the $\mathcal{O}(\hbar^0)$ terms:

$$(S_0')^2 = 2m[E - V(x)]. \tag{3.57}$$

The solution of this equation is

$$S_0(x) = \pm\int_{x_0}^{x} \sqrt{2m[E - V(y)]}dy + c_{1,\pm}. \tag{3.58}$$

Here, $c_{1,\pm}$ is a constant. Next, the $\mathcal{O}(\hbar)$ terms provide the equation

$$2S_0'S_1' = \mathrm{i}S_0'' \to S_1'(x) = \frac{\mathrm{i}}{2}\frac{S_0''}{S_0'}. \tag{3.59}$$

Integration gives:

$$S_1(x) = \frac{\mathrm{i}}{2}\ln S_0'(x) + c_{2,\pm} \tag{3.60}$$

where $c_{2,\pm}$ is a new integration constant. Since we now have identified S_0 and S_1, we find that

$$\psi(x) = \mathrm{e}^{\mathrm{i}S(x)/\hbar} \simeq \mathrm{e}^{\mathrm{i}(S_0+\hbar S_1)/\hbar}, \tag{3.61}$$

and by renaming the constants to $A_\pm \equiv \mathrm{e}^{\mathrm{i}c_{1,\pm}}\mathrm{e}^{\mathrm{i}c_{2,\pm}}$ we thus obtain

$$\psi(x) = \frac{A_\pm}{[E - V(x)]^{1/4}} \mathrm{e}^{\pm\frac{\mathrm{i}}{\hbar}\int_{x_0}^{x} \sqrt{2m[E-V(y)]}dy}. \tag{3.62}$$

This is the WKB approximation for the solution $\psi(x)$. Note that if $V(x) > E$ (classically forbidden area), $\psi(x)$ exponentially increases or decreases since the exponent becomes purely real:

$$\psi(x) = \frac{B_\pm}{[V(x) - E]^{1/4}} \mathrm{e}^{\pm\frac{\mathrm{i}}{\hbar}\int_{x_0}^{x} \sqrt{2m[V(y)-E]}dy}, \tag{3.63}$$

where we absorbed some numerical constants $[(-1)^{1/4}]$ into $A_\pm$ and renamed it to $B_\pm$. For a bound state, one of the coefficients $B_\pm$ thus has to be zero in order to prevent $\psi(x)$ from diverging.

When is WKB valid?
The premise of our approach is that the term $\mathrm{i}\hbar S''$ is small compared to $2m[E - V(x)]$ due to a slowly varying potential. Let $p^2 = 2m[E - V(x)]$. In effect, we demand that

$$|\mathrm{i}\hbar S''| \ll |p^2|. \tag{3.64}$$

If S'' is small, it means that we may approximate

$$(S')^2 - 2m[E - V(x)] - \mathrm{i}\hbar S'' \simeq (S')^2 - p^2 = 0. \tag{3.65}$$

Therefore, $S' = p \to S'' = p'$. This gives us

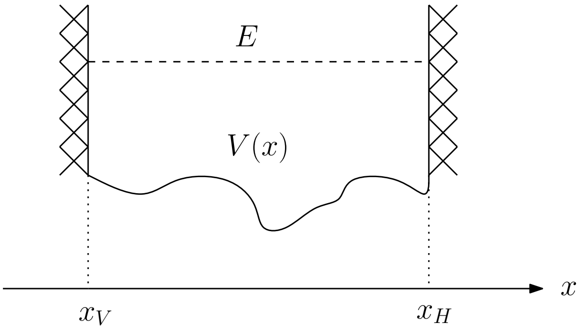
$$|\mathrm{i}\hbar\, dp/dx| \ll |p^2| \to \left|\hbar \frac{d}{dx}\left(\frac{1}{p}\right)\right| \ll 1. \tag{3.66}$$

Since $\hbar/p = \lambda$ is the wavelength of the particle, this means that

$$\boxed{\left|\frac{d\lambda}{dx}\right| \ll 1.}$$

Physically, this means that the change in wavelength λ over a distance λ should be small compared to λ itself in order for the WKB treatment to be valid, which can be satisfied by a slowly varying potential.

Application #1: quantization with hard walls.
Consider a potential containing two hard walls at $x = x_V$ and $x = x_H$, so that $V(x) = \infty$ for $x < x_V$ and $x > x_H$.



Since the wavefunction is zero outside $x_V < x < x_H$, we must have $\psi(x_V) = \psi(x_H) = 0$. We then need a linear combination of the solutions $\psi(x) = \frac{A_\pm}{[E-V(x)]^{1/4}} \mathrm{e}^{\pm\frac{\mathrm{i}}{\hbar}\int_{x_0}^{x} \sqrt{2m[E-V(y)]}dy}$ which vanishes at those points. One combination that satisfies this is

$$\psi(x) = A[E - V(x)]^{-1/4} \sin[(1/\hbar)\int_{x_V}^{x} \sqrt{2m[E - V(y)]}dy] \tag{3.67}$$

if we demand that

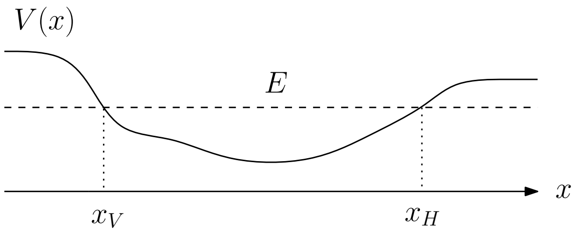
$$\frac{1}{\hbar}\int_{x_V}^{x_H} \sqrt{2m[E - V(y)]}dy = n\pi,\ n = \text{integer}. \tag{3.68}$$

This is effectively a quantization condition for the energy, which may be written as (using $1/\hbar = 2\pi/h$)

$$2\int_{x_V}^{x_H} \sqrt{2m[E_n - V(y)]}dy = nh. \tag{3.69}$$

In the simple limit that $V(y) = 0$, we obtain $2\sqrt{2mE_n}(x_H - x_V) = nh$, which is an exact result.

Application #2: quantization with continuous potential.
Consider a continuously varying potential with a minimum.



We should then expect to have oscillating WKB-solutions for $x_V < x < x_H$, but decaying solutions for $x < x_V$ and $x > x_H$ since those areas are classically forbidden. A problem nevertheless arises at the points where $E = V(x)$ since the WKB-solution diverges there due to the factor $[E - V(x)]^{-1/4}$. The challenge is then: how do we connect the inner solutions $(x_V < x < x_H)$

$$\psi(x) = [E - V(x)]^{-1/4}[A_+ \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\int_{x_V}^{x} \sqrt{2m[E-V(y)]}dy} + A_- \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\int_{x_V}^{x} \sqrt{2m[E-V(y)]}dy}] \tag{3.70}$$

with the outer solutions

$$\begin{aligned}\psi(x) &= B_-[V(x) - E]^{-1/4}\mathrm{e}^{\frac{1}{\hbar}\int_{x_V}^{x} \sqrt{2m[V(y)-E]}dy},\ x < x_V,\\ \psi(x) &= B_+[V(x) - E]^{-1/4}\mathrm{e}^{-\frac{1}{\hbar}\int_{x_H}^{x} \sqrt{2m[V(y)-E]}dy},\ x > x_H.\end{aligned} \tag{3.71}$$

Our strategy will be to treat the areas close to $x = x_V$ and $x = x_H$ exactly, since the potential can then be approximated as linear (via *e.g.* a Taylor expansion), and then use this exact solution to connect the inner and outer solutions. Close to the right turning point $x = x_H$, we have $V(x) - E \simeq c(x - x_H)$ where $c > 0$. The SE for this linear potential reads

$$\frac{d^2\psi}{dx^2} - \frac{2mc}{\hbar^2}(x - x_H)\psi = 0 \tag{3.72}$$

and is equivalent to the previously mentioned Airy's differential equation

$$\frac{d^2\psi}{d\xi^2} - \xi\psi = 0 \tag{3.73}$$

where we introduced $\xi = \left(\frac{2mc}{\hbar^2}\right)^{1/3}(x - x_H)$. We have previously looked at the solutions Ai(x) and Bi(x) and the non-divergent solution Ai(ξ) has the following asymptic behavior:

$$\text{Ai}(\xi) = \begin{cases} \frac{1}{2\sqrt{\pi}}\xi^{-1/4}\text{e}^{-2\xi^{3/2}/3} & \text{for large positive } \xi \\ \frac{1}{\sqrt{\pi}}(-\xi)^{-1/4}\cos[\frac{2}{3}(-\xi)^{3/2} - \frac{\pi}{4}] & \text{for large negative } \xi. \end{cases} \tag{3.74}$$

If we use the linear potential $c(x - x_H)$ in the WKB wavefunction, then for $x < x_H$ we obtain

$$\frac{1}{\hbar}\int_x^{x_H} \sqrt{2m[E - V(y)]}dy = \frac{2}{3}(2mc/\hbar^2)^{1/2}(x_H - x)^{3/2} = \frac{2}{3}(-\xi)^{3/2}. \tag{3.75}$$

We then see that we can write the asymptotic Ai(ξ) function as

$$\text{Ai}(\xi) = \frac{1}{\sqrt{\pi}}\Big[-\Big(\frac{2mc}{\hbar^2}\Big)^{1/3}(x - x_H)\Big]^{-1/4}\cos\Big[\frac{1}{\hbar}\int_x^{x_H}\sqrt{2m[E - V(y)]}dy - \frac{\pi}{4}\Big] \tag{3.76}$$

for large negative ξ. Since $E - V(x) \propto (x - x_H)$, we see that

$$\text{Ai}(\xi) \propto [E - V(x)]^{-1/4}\cos\Big[\frac{1}{\hbar}\int_x^{x_H}\sqrt{2m[E - V(y)]}dy - \frac{\pi}{4}\Big] \tag{3.77}$$

and this is precisely the WKB wavefunction for suitably chosen coefficients $A_\pm$. In other words, by choosing $A_\pm$ so that the WKB wavefunction becomes the asymptotic part of Ai(ξ), we may then connect the inner wavefunction to the outer one for $x > x_H$. Performing the same procedure at the left interface gives us

$$\psi(x) \propto [E - V(x)]^{-1/4}\cos\Big[\frac{1}{\hbar}\int_{x_V}^{x}\sqrt{2m[E - V(y)]}dy - \frac{\pi}{4}\Big]. \tag{3.78}$$

We now have two expressions for the inner wavefunction which should be equal for consistency. Using that $\int_{x_V}^{x} = \int_{x_V}^{x_H} - \int_x^{x_H}$, we can write

$$\psi(x) \propto [E - V(x)]^{-1/4}\cos\Big[\frac{1}{\hbar}\int_x^{x_H}\sqrt{2m[E - V(y)]}dy - \frac{1}{\hbar}\int_{x_V}^{x_H}\sqrt{2m[E - V(y)]}dy + \frac{\pi}{4}\Big]. \tag{3.79}$$

For the two wavefunctions to be equal, we thus obtain the criterion that

$$\frac{1}{\hbar}\int_{x_V}^{x_H}\sqrt{2m[E - V(y)]}dy = n\pi - \frac{\pi}{2}. \tag{3.80}$$

Note that we have used here that the wavefunctions only need to be equal up to an overall sign ± 1 since this sign can be taken care of by the normalization factor. In the above equation, n is an integer. Therefore, it can be rewritten as

$$2\int_{x_V}^{x_H}\sqrt{2m[E - V(y)]}dy = (n - \frac{1}{2})h,\ n = 1, 2, 3, \dots \tag{3.81}$$

The energies E satisfying this equation then determines the energy eigenvalues $E = E_n$. Since the classical energy-momentum relation is $E = p^2/2m + V$, we can write the above result as

$$\oint p(x)dx = (n + \frac{1}{2})h,\ n = 0, 1, 2, \dots \tag{3.82}$$

where the integral is taken over one period of the classical motion (starting and ending up with the same momentum). This is the *Bohr-Sommerfeld quantization condition*.The quantization condition Eq. (3.81) gives better results the larger n is, but decent results may also be obtained for lower-lying levels n as well. For a harmonic oscillator, Eq. (3.81) in fact gives the exact eigenvalues for all n.

We previously treated the case with two hard walls. If the potential instead has one hard wall, *e.g.* at $x = x_V$, the wavefunction must vanish at $x = x_V$. From our expression for the inner wavefunction obtained from the asymptotic behavior at $x = x_H$, we see that the quantization condition becomes: $2\int_{x_V}^{x_H}\sqrt{2m[E-V(y)]}dy = (n-\frac{1}{4})h$. It is then possible to summarize our WKB results for the energy eigenvalues in the presence of hard walls as follows:

- **0 hard walls:** $2\int_{x_V}^{x_H}\sqrt{2m[E-V(y)]}dy = (n-\frac{1}{2})h$.
- **1 hard walls:** $2\int_{x_V}^{x_H}\sqrt{2m[E-V(y)]}dy = (n-\frac{1}{4})h$.
- **2 hard walls:** $2\int_{x_V}^{x_H}\sqrt{2m[E-V(y)]}dy = (n-0)h$.

Example 5. Triangular well. Let us apply the WKB method to the triangular well problem to see how well it approximates the eigenvalues. We have $V(x) = \infty$ for $z < 0$ and $V(z) = Fz$ for $z \geq 0$. This problem thus has one hard wall and to use the quantization condition we have to set $z_V = 0$ and $z_H = E/F$, since z_H was assumed to be located at the classical turning point. We get:

$$2\int_0^{E/F}\sqrt{2m(E-Fz)}dz = (n-\frac{1}{4})h, \tag{3.83}$$

which solving for E provides $E = E_n = \left[\frac{3}{2}\pi\left(n-\frac{1}{4}\right)\right]^{2/3}\left(\frac{\hbar^2}{2m}\right)^{1/3}F^{2/3}$. The numerical coefficients for $n = 1, 2, 3$ are respectively 0.8%, 0.15%, and 0.08% off the exact results!

IV. TIME-DEPENDENT APPROXIMATIVE METHODS

Learning goals. After reading this chapter, the student should:

- Know the fundamental idea behind time-dependent perturbation theory, when it is valid, and be able to mathematically outline how to apply it on a quantum mechanical problem.
- Know the fundamental idea behind the sudden approximation, when it is valid, and be able to mathematically outline how to apply it on a quantum mechanical problem.

So far, we have studied weak or slowly (spatially) varying perturbations of a QM system. Now, we take the step to time dependent perturbations. Important applications for such a framework include EM radiation, spectroscopy, and laser physics.

A. Perturbation theory

We start by considering the case of a weak perturbation (in magnitude) and set out to derive the differential equations governing the state coefficients. Let $\hat{V}(\boldsymbol{r},t)$ be a weak time dependent perturbation:

$$\hat{H}(\boldsymbol{r},t) = \hat{H}^0(\boldsymbol{r}) + \hat{V}(\boldsymbol{r},t). \tag{4.1}$$

Assume that the stationary states $\Psi_n^0(\boldsymbol{r},t) = \psi_n(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}E_n t/\hbar}$ for the unperturbed system $\hat{H}^0(\boldsymbol{r})$ are known. We thus have $\hat{H}^0(\boldsymbol{r})\psi_n = E_n\psi_n$. The time evolution of the time dependent, non-stationary states Ψ are governed by

$$\mathrm{i}\hbar\partial_t\Psi = \hat{H}\Psi. \tag{4.2}$$

We are not able to solve this in its exact form, and thus look for a perturbation method valid for weak $\hat{V}$. Since the eigenstates for the unperturbed system is, as usual, assumed to be a complete and orthonormal set, we may expand

$$\Psi(\boldsymbol{r},t) = \sum_k a_k(t)\psi_k(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}E_k t/\hbar}. \tag{4.3}$$

Note that the coefficients $\{a_k\}$ have to be time dependent. Due to the normalization of $\Psi(\boldsymbol{r},t)$, we obtain

$$\sum_k |a_k(t)|^2 = 1. \tag{4.4}$$

Inserting the expansion Eq. (4.3) into the time dependent SE, we obtain

$$\sum_k \mathrm{i}\hbar\Big(\frac{da_k}{dt} - \frac{\mathrm{i}}{\hbar}E_k a_k\Big)\psi_k(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}E_k t/\hbar} = \sum_k a_k[\hat{H}^0 + \hat{V}(\boldsymbol{r},t)]\psi_k(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}E_k t/\hbar}. \tag{4.5}$$

We know use that $\hat{H}^0\psi_k = E_k\psi_k$ to cancel two terms in the above equations and then multiply it with $[\psi_n(\boldsymbol{r})]^*$ and integrate over space, in order to obtain

$$\mathrm{i}\hbar\frac{da_n}{dt}\mathrm{e}^{-\mathrm{i}E_n t/\hbar} = \sum_k V_{nk}(t)\mathrm{e}^{-\mathrm{i}E_k t/\hbar}. \tag{4.6}$$

We here defined

$$V_{nk}(t) \equiv \int [\psi_n(\boldsymbol{r})]^*\hat{V}(\boldsymbol{r},t)\psi_k(\boldsymbol{r})d\boldsymbol{r} = \langle n|\hat{V}|k\rangle. \tag{4.7}$$

This is a known quantity since it can be computed from the known ψ_n and $\hat{V}$. With the short-hand notation $\omega_{nk} \equiv (E_n - E_k)/\hbar$, we may then write the result as

$$\frac{da_n(t)}{dt} = \frac{1}{\mathrm{i}\hbar}\sum_k V_{nk}\mathrm{e}^{\mathrm{i}\omega_{nk}(t)}a_k(t),\ n = 1,2,3,\ldots \tag{4.8}$$

Note that so far, we have not actually made any approximations: this coupled set of equations for the state coefficients is fully equivalent to the SE.

If we now, however, do assume that $\hat{V}$ is weak, a_k will only have a weak time dependence and we can approximate the solution by neglecting the time dependence of $a_k(t)$ on the r.h.s. of the equation. Doing so and integrating with respect to t gives:

$$a_n(t) = a_n(t_0) + \frac{1}{\mathrm{i}\hbar}\sum_k a_k(t_0)\int_{t_0}^{t} V_{nk}(\tau)\mathrm{e}^{\mathrm{i}\omega_{nk}\tau}d\tau. \tag{4.9}$$

If the system starts out in state b at $t = t_0$, then $a_k(t_0) = \delta_{kb}$ and we obtain

$$a_s(t) = \frac{1}{\mathrm{i}\hbar}\int_{t_0}^{t} V_{sb}(\tau)\mathrm{e}^{\mathrm{i}\omega_{sb}\tau}d\tau \; (s \neq b). \tag{4.10}$$

This is a key result because it tells us that the probability that the system at a time t has made a transition from state b to s is $P_{b\to s}(t) = |a_s(t)|^2$. Note that $P_{b\to b} = 1 - \sum_{s\neq b} P_{b\to s}$.

Detailed balance.
Let us compare the probability for the transitions $b \to s$ and $s \to b$. The first one was calculated above. The second one is

$$a_{s\to b}(t) = \frac{1}{\mathrm{i}\hbar}\int_{t_0}^{t} V_{bs}(\tau)\mathrm{e}^{\mathrm{i}\omega_{bs}\tau}d\tau. \tag{4.11}$$

Since $\omega_{bs} = (E_b - E_s)/\hbar = -\omega_{sb}$ and $V_{bs} = \langle b|\hat{V}|s\rangle = \langle s|\hat{V}|b\rangle^* = V_{sb}^*$, the amplitudes satisfy $a_{b\to s}(t) = -a_{s\to b}^*(t)$. Taking $|\ldots|^2$, we see that

$$P_{s\to b}(t) = P_{b\to s}(t). \tag{4.12}$$

In other words, to first order in time dependent perturbation theory, the probability for a transition is equal to the probability for the opposite transition. This result is known as detailed balance.

Transient perturbations.
Assume that we are dealing with a perturbation that is transient, such as a charged particle passing by an atom and exciting the electrons in the atom. This is actually the dominant mechanism that causes deceleration of an individual charged particle injected into a material. Since the coefficients $\{a_s\}$ stop changing after the perturbation has ceased, we may set $t = \infty$ and use $t_0 = -\infty$ as the initial time. The transition probability from state b to s then takes the form:

$$P_{b\to s} = \Big|\frac{1}{\hbar}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\omega_{sb}\tau}V_{sb}(\tau)d\tau\Big|^2. \tag{4.13}$$

In the special case where $\hat{V}$ varies slowly in time compared to the period ω_{sb}^{-1}, the integrand oscillates rapidly around zero and the integral become very small. If instead the perturbation varies in the same way as the "eigen-frequency" ω_{sb} of the system, a *resonance* can occur which strongly influences the system. We now proceed to consider such a scenario.

B. Harmonic perturbations

An important special case is when the perturbation varies harmonically:

$$\hat{V}(\boldsymbol{r}, t) = V_+(\boldsymbol{r})\mathrm{e}^{\mathrm{i}\omega t} + V_-(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\omega t}. \tag{4.14}$$

The interaction between an atomic system and a radiation field in the form of EM waves has this form. The limit $\omega \to 0$ corresponds to a constant perturbation. In order for $\hat{V}$ to be Hermitian, we need $V_+^* = V_-$. Inserting this $\hat{V}$ into our result for the transition coefficients, we obtain:

$$\begin{aligned} a_{b\to s}(t) &= \frac{1}{\mathrm{i}\hbar}(V_+)_{sb}\int_0^t \mathrm{e}^{\mathrm{i}(\omega_{sb}+\omega)\tau}d\tau + \frac{1}{\mathrm{i}\hbar}(V_-)_{sb}\int_0^t \mathrm{e}^{\mathrm{i}(\omega_{sb}-\omega)\tau}d\tau \\ &= (V_+)_{sb}\frac{1-\mathrm{e}^{\mathrm{i}(\omega_{sb}+\omega)t}}{\hbar\omega_{sb}+\omega\hbar} + (V_-)_{sb}\frac{1-\mathrm{e}^{\mathrm{i}(\omega_{sb}-\omega)t}}{\hbar\omega_{sb}-\omega\hbar}. \end{aligned} \tag{4.15}$$

We have set $t_0 = 0$ as the reference point. To obtain the transition probability we need $|a_s|^2$. This gives $|\ldots|^2$ of the individual terms in Eq. (4.15) and a cross-term. Consider the last term $\propto V_-$ which after $|\ldots|^2$ gives:

$$4|(V_-)_{sb}|^2 \frac{\sin^2[(E_s - E_b - \hbar\omega)t/2\hbar]}{(E_s - E_b - \hbar\omega)^2}. \tag{4.16}$$

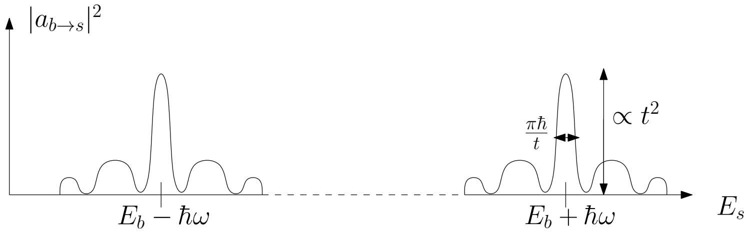
This contribution has a peak (with a height $\propto t^2$) at the energy $E_s = E_b + \hbar\omega$. The width of the peak, on the other hand, goes like t^{-1}. As a crude approximation, we may then write

$$4|(V_-)_{sb}|^2 \frac{\sin^2[(E_s - E_b - \hbar\omega)t/2\hbar]}{(E_s - E_b - \hbar\omega)^2} \simeq |(V_-)_{sb}|^2 \frac{2\pi t}{\hbar} \delta(E_s - E_b - \hbar\omega). \tag{4.17}$$

The term $\propto V_+$ similarly gives a sharp maximum at $E_s = E_b - \hbar\omega$. The cross-term, however, has no sharp maximum and thus for large times t we have the following transition probability per unit time:

$$\omega_{b\to s} = \frac{|a_{b\to s}(t)|^2}{t} \simeq \frac{2\pi}{\hbar}|(V_-)_{sb}|^2 \delta(E_s - E_b - \hbar\omega) + \frac{2\pi}{\hbar}|(V_+)_{sb}|^2 \delta(E_s - E_b + \hbar\omega). \tag{4.18}$$

A sketch of the true behavior of the $|a_{b\to s}|^2$ would look like this:



The formula for $\omega_{b\to s}$ is useful when the energy spectrum or frequencies are continuous so that $E_b \pm \hbar\omega = E_s$ can indeed be satisfied.

Example 6. Continuous ω. An EM field (*e.g.* visible light, X-rays) with a broad spectum of frequencies causes transitions between discrete atomic or molecular states. The resulting absorption spectrum consists of sharp lines.

Continuous E_s. A laser beam with fixed frequency ω can ionize an atom, causing a transition from a discrete bound state to a state in the continuous spectrum. This is the photoelectric effect.

Continuous E_b and E_s. A typical scattering experiment consists of particles in a beam being perturbed by some target (i.e. potential) and changing direction. Such scattering is a transition between different states in continuous spectra.

Transition to continuum states.
Assume that we start out with a state with fixed E_b and that the perturbation has a specific ω, while the final state s lies in a continuum of final states. Let there be $\rho(E)dE$ energy states in the range $(E, E + dE)$, such that $\rho(E)$ is the density of states (DOS). For instance, in previous QM courses you may have shown that the DOS for a free particle in a volume V_0 is

$$\rho(E) = 2\pi(2m/\hbar^2)^{3/2} V_0 E^{1/2}. \tag{4.19}$$

We may then compute the total transition probability to a state with energy close to E_s. This is obtained by

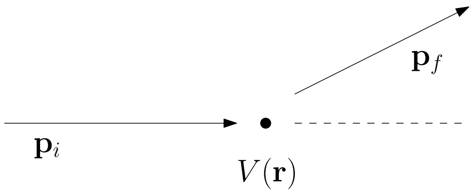
$$\boxed{\omega_{b\to s} = \frac{2\pi}{\hbar}|V_{sb}|^2 \rho(E_s).}$$

The formula expresses that the transition rate ω increases both with the "overlap" $|V_{sb}|$ element between the states and the amount of available states $\rho(E_s)$. This is known as the *golden rule*. If one is interested in only a subset of the states with energy E_s, such as particles moving in a certain direction which a detector can pick up, one simply uses the DOS for that subset. For a free particle, it would be the fraction $d\Omega/4\pi$ of the total DOS:

$$\rho = 2\pi(2m/\hbar^2)^{3/2} V_0 E^{1/2} \frac{d\Omega}{4\pi} = V_0 \frac{m p_f}{h^3} d\Omega, \tag{4.20}$$

where $E = p_f^2/2m$.

We can apply this to a scattering scenario, where a scattering potential $V(\boldsymbol{r})$ acts as a perturbation on a particle-beam $\psi(\boldsymbol{r}) = \frac{1}{\sqrt{V_0}} \mathrm{e}^{\mathrm{i}\boldsymbol{p}_i \cdot \boldsymbol{r}/\hbar}$. The aim is to find the probability per unit time for a transition to the final state $\psi_f(\boldsymbol{r}) = \frac{1}{\sqrt{V_0}} \mathrm{e}^{\mathrm{i}\boldsymbol{p}_f \cdot \boldsymbol{r}/\hbar}$. Here, V_0 is the volume under consideration. We may treat this process as stationary (corresponding to $\omega = 0$) and the energy before and after is thus the same: $|\boldsymbol{p}_f| = |\boldsymbol{p}_i| \equiv p$.



To obtain $\omega_{b\to s}$ via the golden rule, we need the matrix-element:

$$V_{fi} = \frac{1}{V_0} \int \mathrm{e}^{\mathrm{i}(\boldsymbol{p}_i - \boldsymbol{p}_f)\cdot \boldsymbol{r}/\hbar} V(\boldsymbol{r}) d\boldsymbol{r}. \tag{4.21}$$

It follows that

$$\omega_{i\to f} = \frac{2\pi}{\hbar} \frac{1}{V_0^2} \left| \int \mathrm{e}^{\mathrm{i}(\boldsymbol{p}_i - \boldsymbol{p}_f)\cdot \boldsymbol{r}/\hbar} V(\boldsymbol{r}) d\boldsymbol{r} \right|^2 \times V_0 \frac{mp}{h^3} d\Omega. \tag{4.22}$$

A common way to measure scattering is the scattering cross section $d\sigma$:

$$d\sigma = \frac{\text{number of particles scattered into } d\Omega \text{ per unit time}}{\text{incident particle intensity}}. \tag{4.23}$$

Quantitatively, the nominator is $\omega_{i\to f}$ and the incoming particle intensity is the product of the particle density $|\psi_i|^2 = 1/V_0$ and the velocity p/m so that

$$d\sigma = \omega_{i\to f}\frac{mV_0}{p}. \tag{4.24}$$

With our expression for $\omega_{i\to f}$, we end up with

$$\boxed{\frac{d\sigma}{d\Omega} = \left|\frac{m}{2\pi\hbar^2}\int V(\boldsymbol{r})\mathrm{e}^{\mathrm{i}(\boldsymbol{p}_i-\boldsymbol{p}_f)\cdot\boldsymbol{r}/\hbar}d\boldsymbol{r}\right|^2.}$$

Later, we will regain this result using a different method. The above formula is the so-called *Born approximation* for the scattering cross section. In chapter 13, we will also examine the range of validity for this result.

C. Sudden approximation

Let us now consider a scenario where the magnitude of the perturbation is not necessarily weak, but where the disturbance is switched on very abruptly. The simplest scenario where one can envision this is where H changes abruptly from H_0 to H_1 at $t = 0$, where H_0 and H_1 are both time independent in themselves. Thus, we have

$$t < 0: \; H_0\psi_k^0 = E_k^0\psi_k^0 \tag{4.25}$$

where ψ_k^0 are orthonormal and form a complete set, which is not necessarily discrete. Moreover,

$$t > 0: \; H_1\phi_n^1 = E_n^1\phi_n^1 \tag{4.26}$$

where $\{\phi_n^1\}$ are also orthonormal and complete. The general solution of the time dependent SE is then:

$$\begin{aligned} t < 0 : \Psi(t) &= \sum_k c_k^0\psi_k^0\mathrm{e}^{-\mathrm{i}E_k^0t/\hbar}, \\ t > 0 : \Psi(t) &= \sum_n d_n^1\psi_n^1\mathrm{e}^{-\mathrm{i}E_n^1t/\hbar}. \end{aligned} \tag{4.27}$$

Assume that $\Psi(t)$ is normalized to unity, so that c_k^0 and d_n^0 are the usual probability coefficients for finding the system in state ψ_k^0 and ϕ_n^1 at $t < 0$ and $t > 0$. Now, since the time dependent SE is first order in the time coordinate, it means that $\Psi(t)$ must be a continuous function of t. Thus, at $t = 0$:

$$\sum_k c_k^0\psi_k^0 = \sum_n d_n^1\phi_n^1. \tag{4.28}$$

Take the scalar product with ϕ_n^1:

$$d_n^1 = \sum_k c_k^0\langle\phi_n^1|\psi_k^0\rangle. \tag{4.29}$$

We now have a way to obtain the probability coefficients after the sudden change at $t = 0$, given that $\{c_k^0\}$ are known. In practice, the change from H_0 to H_1 will take place over a short time interval τ rather than being instantaneous. The simplest way to approximate this scenario is to use Eq. (4.29), but how large can τ while Eq. (4.29) remains useful?

We derive a simple criterion of validity. Let:

$$H = \begin{cases} H_0 \text{ for } t < 0 \\ H_i \text{ for } 0 < t < \tau \\ H_1 \text{ for } t > \tau \end{cases} \tag{4.30}$$

where H_i is the time independent Hamiltonian during the intermediate period τ. If $\{\chi_l^i\}$ denotes the complete orthonormal set of eigenfunctions of H_i, so that $H_i\chi_l^i = E_l^i\chi_l^i$, then the general solution for the state coefficients $\{d_n^1\}$ determining the state at $t > \tau$ can be found in the same way as above, namely by using continuity of the wavefunction at $t = 0$ and $t = \tau$. It yields:

$$d_n^1 = \sum_k\sum_l c_k^0\langle\phi_n^1|\chi_l^i\rangle\langle\chi_l^i|\psi_k^0\rangle\mathrm{e}^{\mathrm{i}(E_n^1-E_l^i)\tau/\hbar}. \tag{4.31}$$

Compare this with Eq. (4.29) which was obtained for $\tau = 0$, i.e. instantaneous switch from H_0 to H_1 rather than through an intermediate Hamiltonian H_i. If we set $\tau = 0$ in Eq. (4.31), the equations are equivalent as expected. However, if $\tau \neq 0$, the difference arises because of $\mathrm{e}^{\mathrm{i}(E_n^1 - E_l^i)\tau/\hbar}$ not being unity. If the sudden approximation is to be valid, we thus need τ to be small compared to all the inverse energy differences $\hbar/|E_n^1 - E_l^i|$ so that the exponential is close to unity:

$$\tau \ll \hbar/|E_n^1 - E_l^i|. \tag{4.32}$$

An interesting special case of the sudden approximation is when the system initially ($t < 0$) is in a particular stationary state $\psi_a^0 \mathrm{e}^{-\mathrm{i}E_a^0 t/\hbar}$ where ψ_a^0 is an eigenstate of H_0. Then, $c_k^0 = \delta_{ka}$ and the probability amplitude of finding the system in eigenstate ϕ_n^1 of H_1 after the sudden change in the Hamiltonian has occurred is simply $d_n^1 = \langle \phi_n^1 | \psi_a^0 \rangle$.

Example 7. Beta decay of the tritium nucleus. A tritium atom consists of a nuclear ^{3}H (one proton + two neutrons) and one electron. It is unstable and decays into the nucleus ^{3}He (two protons + one neutron):

$$^3\text{H} \to \ ^3\text{He} + e^- + \bar{\nu}_e. \tag{4.33}$$

Assume that the tritium atom is in its ground state before the β-decay of ^{3}H takes place. The question is now: what is the influence of the decay on the atomic electron?

We first note that in the β-decay process above, the electron is emitted from the nucleus with, in most cases, an energy of several keV. This means that its resulting velocity v is much higher than the velocity $v_0 \simeq c/137$ of the atomic electron in the ground state of tritium. If a_0 is the Bohr radius, the emitted electron will leave the atom in a time $\tau \simeq a_0/v$. This is much shorter than the period $T = 2\pi a_0/v_0$ associated with the motion of the atomic electron. Thus, we can justify a scenario where the nuclear charge "seen" by the atomic electron changes instantaneously from Ze to $Z'e$ where $Z = 1$ and $Z' = 2$. The relevant Hamiltonians we have to work with are then:

$$\begin{aligned} H(t<0) = H_0 &= -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r}, \\ H(t>0) = H_1 &= -\frac{\hbar^2}{2m}\nabla^2 - \frac{Z'e^2}{4\pi\epsilon_0 r}. \end{aligned} \tag{4.34}$$

with m being the mass of the atomic electron. We neglected here the recoil effect on the nucleus, since its mass $M \gg m$. The eigenfunctions of H_0 and H_1 are hydrogenic wavefunctions and thus known. Since the tritium atom is assumed to initially be in its ground state (quantum nunbers $n = 1, l = 0, m = 0$), the probability coefficients $d_{n'l'm'}^1$ of finding the atomic electron in a discrete eigenstate $(n'l'm')$ of H_1 at $t > 0$ is:

$$d_{n'l'm'}^1 = \langle \psi_{n'l'm'}^{(Z'=2)} | \psi_{100}^{(Z=1)} \rangle = \int \left(\psi_{n'l'm'}^{(Z'=2)}(\boldsymbol{r}) \right)^* \psi_{100}^{(Z=1)}(\boldsymbol{r}) d\boldsymbol{r} \tag{4.35}$$

where $\psi_{nlm}^{(Z)}(\boldsymbol{r})$ is a hydrogenic wavefunction with atomic number Z. We know that $\psi_{nlm}^{(Z)}(\boldsymbol{r}) = R_{nl}^{(Z)}(\boldsymbol{r}) Y_{lm}(\theta, \phi)$, and from the orthonormality properties of Y_{lm} one may verify that the only non-vanishing probability coefficients $d_{n'l'm'}^1$ are those belonging to the s-states ($l' = m' = 0$):

$$d_{n'00}^1 = \int_0^\infty R_{n'0}^{(Z'=2)}(r) R_{10}^{(Z=1)}(r) r^2 dr. \tag{4.36}$$

For the particular case $n' = 1$, we obtain

$$d_{100}^1 = 2^{7/2} a_0^{-3} \int_0^\infty dr\, r^2 \mathrm{e}^{-3r/a_0} = \frac{16\sqrt{2}}{27}. \tag{4.37}$$

Hence, the probability that the ^{3}He ion is found in its ground state is $P_{100}^1 = |d_{100}^1|^2 \simeq 0.702$. The total probability for the ion to be either excited or even ionized is then $1 - P_{100}^1 \simeq 0.298$.

V. ADIABATIC APPROXIMATION AND THE BERRY PHASE

Learning goals. After reading this chapter, the student should:

- Know the fundamental idea behind the adiabatic approximation, when it is valid, and be able to mathematically outline how to apply it on a quantum mechanical problem.
- Be able to explain what the Berry phase is and in which scenario it is of relevance. The student should also be able to give concrete examples of systems where the Berry phase plays an important role.

The lecture notes forming the basis for this chapter follow roughly the same structure as the corresponding chapters in "Quantum Mechanics" by Bransden & Joachain.

A. The adiabatic approximation

The perturbation method we have initially considered was based on the assumption that the magnitude of the time dependent part of H has been small. We now present a new approximation where the key parameter is the *rate of change* of H. Start by assuming that H varies very slowly with time, i.e. the completely opposite scenario of the sudden approximation. One should then expect that the approximate solution of $\mathrm{i}\hbar\partial_t\Psi = H(t)\Psi$ can be obtained in terms of the eigenfunctions $\psi_k(t)$ of the "instantaneous" Hamiltonian $H(t)$ so that

$$H(t)\psi_k(t) = E_k(t)\psi(t) \tag{5.1}$$

at any given time t. Physically, what we are stating here is that if $H(t)$ changes very slowly, a system which at $t = t_0$ is in a discrete non-degenerate state $\psi_a(t_0)$ with energy $E_a(t_0)$ is very likely to be in the state $\psi_a(t)$ with energy $E_a(t)$ at a later time t, i.e. without making any transition. We now proceed to prove this *adiabatic theorem*, using the method of Born & Fock from 1928.

In fact, the original formulation of this theorem was:

> A physical system remains in its instantaneous eigenstate if a given perturbation is acting on it slowly enough and if there is a gap between the eigenvalue and the rest of the Hamiltonian's spectrum.

We can think of this physically as follows: if a QM system is exposed to a slowly changing perturbation, the system has time to adapt to the perturbation. In contrast, if the perturbation occurs very rapidly there is not sufficient time for the system to adapt, so that the probability density remains unaltered.

Assume that Ψ is known at $t = t_0$. Now, expand for $t \geq t_0$ in the instantaneous eigenfunctions $\psi_k(t)$:

$$\Psi = \sum_k c_k(t)\psi_k(t)\text{exp}\Big[-(\text{i}/\hbar)\int_{t_0}^t E_k(t')dt'\Big]. \tag{5.2}$$

We assume that $\{\psi_k\}$ form an orthonormal and complete set, as usual. The energies $E_k(t)$ are non-degenerate and form a discrete spectrum. Note that the "energy levels" is just a formal name, since energy is not strictly speaking conserved for a time dependent H. Inserting the above expansion into the time dependent SE provides:

$$\text{i}\hbar\sum_k\Big(\dot{c_k}\psi_k + c_k\partial_t\psi_k - (\text{i}/\hbar)c_k\psi_k E_k\Big)\text{exp}\Big[-(\text{i}/\hbar)\int_{t_0}^t E_k(t')dt'\Big] = H(t)\sum_k c_k\psi_k\text{exp}\Big[-(\text{i}/\hbar)\int_{t_0}^t E_k(t')dt'\Big]. \tag{5.3}$$

There is a cancellation of the last term on the l.h.s. by using Eq. (5.1). Now, do the following:

- Multiply with $\psi_b^*(t)$ (which is part of the set $\{\psi_k(t)\}$).
- Integrate over the coordinates of the system.
- Use that $\langle\psi_b|\psi_k\rangle = \delta_{bk}$.

This gives:

$$\dot{c}_b(t) = -\sum_k c_k(t)\text{exp}\Big\{\frac{\text{i}}{\hbar}\int_{t_0}^t[E_b(t') - E_k(t')]dt'\Big\}\langle\psi_b|\partial_t\psi_k\rangle. \tag{5.4}$$

This is a set of coupled first order differential equations for all the coefficients $c_k(t)$. The diagonal terms can be removed as follows. Consider first $\alpha_k(t) = \langle\psi_k|\partial_t\psi_k\rangle$. Use the normalization $\langle\psi_k(t)|\psi_k(t)\rangle = 1$ and differentiate it with respect to time:

$$\langle\partial_t\psi_k|\psi_k\rangle + \langle\psi_k|\partial_t\psi_k\rangle = [\alpha_k(t)]^* + \alpha_k(t) = 0. \tag{5.5}$$

Thus, $\alpha_k(t)$ is purely imaginary, so that we may write $\alpha_k(t) = \text{i}\beta_k(t)$ where $\beta \in \Re$. Now, define

$$c_k'(t) = c_k(t)\text{e}^{\text{i}\int_{t_0}^t \beta_k(t')dt'}. \tag{5.6}$$

Differentiating c_b' with respect to time in order to get $\dot{c}_b'$, we obtain:

$$\dot{c}_b' = -\sum_{k\neq b} c_k'(t)\text{exp}\Big\{\frac{\text{i}}{\hbar}\int_{t_0}^t[E_b'(t') - E_k'(t')]dt'\Big\}\langle\psi_b'|\partial_t\psi_k'\rangle, \tag{5.7}$$

where $E_k'(t) = E_k(t) + \hbar\beta_k(t)$. We defined $\psi_k'(t)$ via:

$$c_k(t)\psi_k(t) = c_k(t)\text{e}^{\text{i}\int_{t_0}^t \beta_k(t')dt'}\psi_k(t)\text{e}^{-\text{i}\int_{t_0}^t \beta_k(t')dt'} \equiv c_k'(t)\psi_k'(t). \tag{5.8}$$

If we assume that the phases of the eigenfunctions ψ_k are arbitrary at each instant of time, we can do this change on all ψ_k. Assume from now on that this change has been made and we thus omit the $'$ notation. It is important to note that this assumption is invalid for the case of *cyclic systems*, but we return to this issue later.

Looking at Eq. (5.7) again, we examine $\langle\psi_b|\partial_t\psi_k\rangle$ for $k \neq b$. Differentiate w.r.t. time on the equation $H(t)\psi_k(t) = E_k(t)\psi_k(t)$ and obtain:

$$\partial_t H\psi_k + H\partial_t\psi_k = \partial_t E_k\psi_k + E_k\partial_t\psi_k. \tag{5.9}$$

Using the notation $\langle\psi_a|\psi_b\rangle = \int[\psi_a(\boldsymbol{r})]^*\psi_b(\boldsymbol{r})d\boldsymbol{r}$ and taking the scalar product with ψ_b, this gives:

$$\langle\psi_b|\partial_t H|\psi_k\rangle + \langle\psi_b|H|\partial_t\psi_k\rangle = E_k\langle\psi_b|\partial_t\psi_k\rangle. \tag{5.10}$$

Using that H is Hermitian, we obtain for the second term that

$$\langle\psi_b|H|\partial_t\psi_k\rangle = E_b\langle\psi_b|\partial\psi_k\rangle, \tag{5.11}$$

and plugging this back into Eq. (5.10) gives:

$$\langle\psi_b|\partial_t\psi_k\rangle = -\frac{(\partial_t H)_{bk}}{\hbar\omega_{bk}(t)},\ b \neq k. \tag{5.12}$$

We introduced the notation $(\partial_t H)_{bk} = \langle\psi_b|\partial_t H|\psi_k\rangle$ and $\omega_{bk}(t) = [E_b(t) - E_k(t)]/\hbar$, $b \neq k$. Thus, $\omega_{bk} \neq 0$ always since we assumed that the energy levels were non-degenerate.

If we now use our obtained results and plug them back into our expression for the coupled equations for the $c_k(t)$ coefficients, we obtain (keep in mind that we omit the primes, as explained previously):

$$\dot{c}_b(t) = \sum_{k\neq b}\frac{c_k(t)}{\hbar\omega_{bk}(t)}(\partial_t H)_{bk}\mathrm{e}^{\mathrm{i}\int_{t_0}^t \omega_{bk}(t')dt'}. \tag{5.13}$$

This system of equations then determine the c_b coefficients, which in turn determine the wavefunction via Eq. (5.2). This is a convenient starting point to make approximations, especially when $\partial_t H$ is small (slowly varying Hamiltonian in time). If $\partial_t H = 0$, then the solution is seen to be simply c_b =constant for all b. If $\partial_t H$ is finite, but small, we can try to solve Eq. (5.13) by setting all c_k on the r.h.s. to be constants. Assume that the system initially ($t = t_0$) is in a state a. We substitute the values $c_k = \delta_{ka}$ in the r.h.s. and get

$$\dot{c}_b(t) = \hbar^{-1}\omega_{ba}^{-1}(t)(\partial_t H)_{ba}\mathrm{e}^{\mathrm{i}\int_{t_0}^t \omega_{ba}(t')dt'},\ b \neq a. \tag{5.14}$$

For $b = a$, we get $\dot{c}_a = 0$ in this approximation. Now integrate the above equation with the initial condition $c_b(t \leq t_0) = 0$ ($b \neq a$) and obtain:

$$\boxed{c_b(t) = \hbar^{-1}\int_{t_0}^t dt'\omega_{ba}^{-1}[\partial_{t'}H(t')]_{ba}\exp\Big[\mathrm{i}\int_{t_0}^t \omega_{ba}(t'')dt''\Big],\ (b \neq a).}$$

This is the result for the adiabatic approximation for the probability amplitude $c_b(t)$. We should expect this result to yield a small $|c_b(t)|$ in order to be valid. Thus, $P_{ba}(t) = |c_b(t)|^2$ denotes the transition probability from the initial state a to state b, and we must have $P_{ba}(t) \ll 1$.

A crude estimate is to assume that ω_{ba} and $\partial_t H$ are time independent. We then obtain

$$\begin{aligned} c_b(t) &\simeq (\mathrm{i}\hbar)^{-1}\omega_{ba}^{-2}(\partial_t H)_{ba}(\mathrm{e}^{\mathrm{i}\omega_{ba}(t-t_0)} - 1), \\ P_{ba}(t) &\simeq 4\hbar^{-2}\omega_{ba}^{-4}|(\partial_t H)_{ba}|^2\sin^2[\omega_{ba}(t-t_0)/2]. \end{aligned} \tag{5.15}$$

This probability behaves reasonably as time increases since it merely oscillates, and the upper bound is [since $\sin^2(x) \leq 1$]:

$$P_{ba}(t) \leq \frac{4|(\partial_t H)_{ba}|^2}{\hbar^2\omega_{ba}^4}. \tag{5.16}$$

The adiabatic approximation is thus valid if

$$|(\partial_t H)_{ba}|^2 \ll \frac{\hbar^2\omega_{ba}^4}{4}. \tag{5.17}$$

Example 8. Charged harmonic oscillator in a time dependent electric field. Let us now try out the adiabatic approxmation and consider a charged particle subject to a linear harmonic oscillator potential and a spatially uniform, time dependent electric field $\mathcal{E}(t)$. The Hamiltonian is then:

$$H(t) = -\frac{\hbar^2}{2m}\partial_x^2 + \frac{1}{2}kx^2 - q\mathcal{E}(t)x = \frac{-\hbar^2}{2m}\partial_x^2 + \frac{1}{2}k[x-a(t)]^2 - \frac{1}{2}ka^2(t). \tag{5.18}$$

Here, we defined $a(t) = q\mathcal{E}(t)/k$. We can physically interpret this $H(t)$ as, at a given time t, describing a standard harmonic oscillator with frequency $\omega = \sqrt{k/m}$, but displaced equilibrium position to $x = a(t)$. The term $-\frac{1}{2}ka^2(t)$ is just a constant. The instantaneous energy eigenfunctions are thus obtained as:

$$\psi_n = \Big(\frac{\alpha}{\sqrt{\pi}2^n n!}\Big)^{1/2}\text{exp}[-\alpha^2(x-a)^2/2]H_n[\alpha(x-a)], \tag{5.19}$$

where $\alpha \equiv \sqrt{m\omega/\hbar}$. The corresponding instantaneous energy eigenvalues take the form

$$E_n(t) = (n+1/2)\hbar\omega - ka^2(t)/2,\ n = 0, 1, 2, \dots \tag{5.20}$$

The angular frequencies $\omega_{nn'} = [E_{n'}(t) - E_n(t)]/\hbar = (n'-n)\omega$ are thus independent of time and equal to the unperturbed value. Assume now that

- $\mathcal{E}(t)$ is applied at $t = t_0$ and that it varies slowly.
- The harmonic oscillator is initially in its ground state ($n = 0$).

We want to compute the probability that the system is in an excited state at $t = t_1$. First, note that $\partial_t H = -k\dot{a}x$ with $\dot{a} = (q/k)(d\mathcal{E}/dt)$. To find the transition probabilities, we will need (as derived previously)

$$\langle\psi_b|\partial_t H|\psi_0\rangle = (\partial_t H)_{b0}. \tag{5.21}$$

In effect, we need to compute matrix elements of the type

$$x_{b0} \equiv \langle\psi_b|x|\psi_0\rangle. \tag{5.22}$$

It can be shown that all these matrix elements vanish when $b \neq 1$, while for $b = 1$ we have $x_{10} = \sqrt{\hbar/(2m\omega)}$. From the general expression of $P_{ba}(t)$ derived previously, the only non-vanishing transition probability is $0 \to 1$. Inserting our expression for $\partial_t H$ and $\omega_{10} = \omega$, we get

$$P_{10}(t_1) = |c_1(t_1)|^2 = \frac{q^2}{2m\hbar\omega^3}\left|\int_{t_0}^{t_1}\frac{d\mathcal{E}(t)}{dt}\mathrm{e}^{\mathrm{i}\omega(t-t_0)}dt\right|^2. \tag{5.23}$$

The slower the $\mathcal{E}$-field varies, the smaller the transition probability $0 \to 1$.

B. The Berry phase

When we discussed the adiabatic approximation, it was assumed that the phases of the eigenfunctions $\psi_k(t)$ are arbitrary at each instant of time. This was in fact generally accepted up to 1984 when M. V. Berry showed that:

> In a cyclic system where the Hamiltonian at time t_f is the same as at time t_0, there is a relative change in the phase between $\psi_k(t_0)$ and $\psi_k(t_f)$ which cannot be removed by a phase transformation and thus has observable consequences.

To show this, consider the case where $H(t)$ varies so slowly that the system remains in its initial non-degenerate state with energy $E_a(t)$ and eigenfunction $\psi_a(t)$. According to our previous treatment of such an adiabatic scenario, the approximate solution of $\mathrm{i}\hbar\partial_t\Psi = H(t)\Psi$ is then:

$$\Psi(t) = c_a(t)\psi_a(t)\mathrm{e}^{-(\mathrm{i}/\hbar)\int_{t_0}^{t}E_a(t')dt'}. \tag{5.24}$$

Since $\Psi(t)$ is in the state $\psi_a(t)$ at $t = t_0$, we can set $c_a(t_0) = 1$. Moreover, since $\psi_a(t)$ should also be normalized to unity, we can write generally at $t \neq t_0$ that $c_a(t) = \mathrm{e}^{\mathrm{i}\gamma_a(t)}$ with $\gamma_a(t) \in \Re$ and $\gamma_a(t_0) = 0$. Now, $\mathrm{e}^{-(\mathrm{i}/\hbar)\int_0^t E_a(t')dt'}$ is the usual dynamical phase factor whereas the SE gives us the following equation for $\gamma_a(t)$:

$$\mathrm{i}\dot{\gamma}_a(t)\psi_a(t) = -\partial_t\psi_a(t), \tag{5.25}$$

with the solution

$$\gamma_a(t) = \mathrm{i}\int_{t_0}^{t} \langle\psi_a(t')|\partial_{t'}\psi_a(t')\rangle dt'. \tag{5.26}$$

If the system is cyclic, then $H(t_f) = H(t_0)$ since the Hamiltonian returns to its value at $t = t_0$ at a later time $t = t_f$. This also implies that $E_a(t_f) = E_a(t_0)$ and $\psi_a(t_0) = \psi_a(t_f)$. The *Berry phase* is the accumulated phase change from t_0 to t_f:

$$\boxed{\bar{\gamma}_a \equiv \mathrm{i}\int_{t_0}^{t_f} \langle\psi_a(t')|\partial_{t'}\psi_a(t')\rangle dt'.}$$

This turns out to be a physically observable quantity, thus with experimentally verifiable quantities. Importantly, the Berry-phase is gauge-invariant and a key message in the 1984 paper by Berry was that any gauge-invariant quantity is in principle observable. In fact, let us see what happens if we try to eliminate $\bar{\gamma}_a$ by transforming $\psi_a \to \psi'_a = \psi_a \mathrm{e}^{\mathrm{i}\eta(t)}$. Under this transformation, the Berry phase becomes $\bar{\gamma}'_a$ where:

$$\begin{aligned}\bar{\gamma}'_a &= \mathrm{i}\int_{t_0}^{t_f} \langle \psi'_a(t')|\partial_{t'}\psi'_a(t')\rangle dt' \\ &= \mathrm{i}\int_{t_0}^{t_f} \langle \psi_a(t')|\partial_{t'}\psi_a(t')\rangle dt' - \int_{t_0}^{t_f} \frac{d\eta(t')}{dt'}dt' \\ &= \bar{\gamma}_a - \eta(t_f) + \eta(t_0).\end{aligned} \tag{5.27}$$

Since $\psi_a(t_f) = \psi_a(t_0)$, it follows that $\eta(t_f) - \eta(t_0) = 2\pi n$, $n = 0, 1, 2, \ldots$ Thus, $\mathrm{e}^{\mathrm{i}\bar{\gamma}_a} = \mathrm{e}^{\mathrm{i}\bar{\gamma}'_a}$ cannot be removed by a phase transformation. Strictly speaking, the Berry phase $\bar{\gamma}_a$ is gauge invariant up to an integer multiple of 2π, whereas $\mathrm{e}^{\mathrm{i}\bar{\gamma}_a}$ is absolutely gauge invariant and thus related to physical observables.

The $H(t)$ may be time dependent through a number of parameters, each of which slowly vary with time. A common example: components of an exernal electric or magnetic field which interact with the system. Consider the case where $H(t)$ depends on t via three parameters $R_1(t), R_2(t), R_3(t)$:

$$H(t) = H[R_i(t)],\ i = 1, 2, 3. \tag{5.28}$$

Since $H(t_f) = H(t_0)$ for a cyclic Hamiltonian, we have $R_i(t_f) = R_i(t_0)$. In vector notation, $\boldsymbol{R} = (R_1, R_2, R_3)$, we can then write the Berry phase as

$$\bar{\gamma}_a = \mathrm{i}\oint \langle \psi_a(\boldsymbol{R})|\nabla_{\boldsymbol{R}}\psi_a(\boldsymbol{R})\rangle \cdot d\boldsymbol{R} \tag{5.29}$$

where $\nabla_{\boldsymbol{R}}$ is the gradient in parameter space and the closed integral is taken along the curve $\mathcal{C}$ in parameter space. We define the Berry connection:

$$\boldsymbol{A}(\boldsymbol{r}) \equiv \mathrm{i}\langle \psi_a(\boldsymbol{R})|\nabla_{\boldsymbol{R}}\psi_a(\boldsymbol{R})\rangle. \tag{5.30}$$

Since it depends on the closed curve $\mathcal{C}$, the Berry phase is often called a *geometrical phase*. Such phases arise also in a number of non-adiabatic situations as well - not only the strictly adiabatic context discussed here. In fact, a generalization of Berry's phase is the Aharonov-Anandan phase. Suppose a system evolves according to the SE, but that the change in H is neither adiabatic or cylic. The system can then still exhibit a geometrical phase: all that is needed is a cyclic evolution of the state of the system. Such a cyclic evolution defines a closed path $\mathcal{C}$ in the Hilbert space of the state. Regardless of whether this evolution is adiabatic or not, it leaves the system with a dynamical phase which depends on the Hamiltonian, and a geometrical phase which depends on the path $\mathcal{C}$.

We also remark that by applying Stoke's theorem, we have:

$$\bar{\gamma}_a = \oint_{\mathcal{C}} \boldsymbol{A}(\boldsymbol{R}) \cdot d\boldsymbol{R} = \int\int \boldsymbol{B} \cdot d\boldsymbol{S} \tag{5.31}$$

where S is the surface bound by the closed path $\mathcal{C}$ and

$$\boldsymbol{B} = \nabla_{\boldsymbol{R}} \times \boldsymbol{A}(\boldsymbol{R}) \equiv \text{ Berry curvature.} \tag{5.32}$$

When treating the Aharonov-Bohm effect, we will see a concrete example of the physical consequences of these kind of geometrical phases.

In closing, we comment on whether or not the Berry/geometrical phase is reconcilable with the commonly stated fact that the overall phase of a quantum system is unobservable. *Yes*, because the Berry phase expresses the total phase change acccumulated during a cycle (either a cyclic evolution of the state or the Hamiltonian). We assumed for simplicity in our derivation that we know the phase at $t = t_0$ was $\gamma_a(t = t_0) = 0$, but generally the Berry phase expresses the *phase difference*:

$$\bar{\gamma}_a = \gamma(t = t_f) - \gamma_a(t = t_0) = \mathrm{i}\int_{t_0}^{t_f} \langle \psi_a(t')|\partial_{t'}\psi_a(t')\rangle dt'. \tag{5.33}$$

Now, phase differences are certainly observable, even if the phase at a given time is not.

Example 9. Relative phase in a superposition of states. A quantum system may be in a superposition of states. Then, the relative phase between the state is observable:

$$\psi = \psi_A + \psi_B = |\psi_A|\mathrm{e}^{\mathrm{i}\alpha_A} + |\psi_B|\mathrm{e}^{\mathrm{i}\alpha_B} = (|\psi_A|\mathrm{e}^{\mathrm{i}(\alpha_A - \alpha_B)} + |\psi_B|)\mathrm{e}^{\mathrm{i}\alpha_B}. \tag{5.34}$$

It is clear that $|\psi|^2$ depends on $\Delta\alpha = \alpha_A - \alpha_B$. More generally, consider two paths $\boldsymbol{R}(t)$ and $\boldsymbol{R}'(t)$ with the same end-points: $\boldsymbol{R}(t_0) = \boldsymbol{R}'(t_0)$ and $\boldsymbol{R}(t_f) = \boldsymbol{R}'(t_f)$. If the system now evolves in a superposition of states $|\psi_i[\boldsymbol{R}(t)]\rangle$ and $|\psi_i[\boldsymbol{R}'(t)]\rangle$, then the relative phase of this superposition (analogously to $\Delta\alpha$ above) contains two parts at $t = t_f$:

- The relative dynamical phase.
- The Berry phase: the difference between the Berry connection $\boldsymbol{A}$ integrated along $\boldsymbol{R}$ and $\boldsymbol{A}$ integrated along $\boldsymbol{R}'$. In effect, it is the circular integral $\oint_{\mathcal{C}} \boldsymbol{A}(\boldsymbol{r}) \cdot d\boldsymbol{r}$ where $\mathcal{C}$ is the closed path comprised of the paths $\boldsymbol{R}$ and $\boldsymbol{R}'$.

VI. QUANTUM MECHANICAL SCATTERING THEORY

Learning goals. After reading this chapter, the student should:

- Be able to explain what the scattering cross section is and what it provides information about.
- Understand how scattering can be formulated as a stationary problem in quantum mechanics and set up the corresponding asymptotic wavefunction, as well as how to modify this for scattering of identical particles.
- Know the underlying idea behind the Born approximation and the method of partial waves, and explain when these two frameworks can be used.
- Know what the optical theorem states physically and the principle from which it is derived.

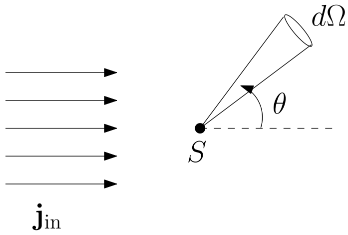
A. Intro to scattering cross section

Particles that are incident toward a scattering center - *e.g.* a different particle - will in general be deflected due to the interaction. The distribution of angles of deflection will depend on the details of the setup. Experimentally measuring this distribution will provide us with information about the type of interaction that is in play. We distinguish between

- **Elastic scattering:** the kinetic energy of the scattered particles is preserved.
- **Inelastic scattering**: kinetic energy is not conserved, *e.g.* due to a photon taking off with part of the energy.

We will consider elastic scattering with the two same particles before and after. For a potential $V(\boldsymbol{r}_1 - \boldsymbol{r}_2)$, we know that this can be reduced to an effective one-body problem where only the relative motion of the particles matter. We shall initially consider the scattering problem in the corresponding center-of-mass (CM) frame and later see how the results are expressed in the lab-frame.

Consider the following idealized model.



A uniform flux of particles with density $\boldsymbol{j}_{\text{in}}$ is incident on a scattering center S. A detector counts particles scattered into solid angle $d\Omega = \sin\theta d\theta d\phi$ enclosing the direction (θ, ϕ). The incident axis is $\theta = 0$. We have previously (chapter 4) defined

$$\frac{d\sigma}{d\Omega} = \frac{\text{\# particles scattered into } d\Omega \text{ per unit time}}{d\Omega \cdot j_{\text{in}}}. \tag{6.1}$$

Since $j_{\text{in}} = |\boldsymbol{j}_{\text{in}}|$ is the number of particles incident per time and area, inspection shows that $d\sigma/d\Omega$ has dimension area. The total scattering cross section is obtained as:

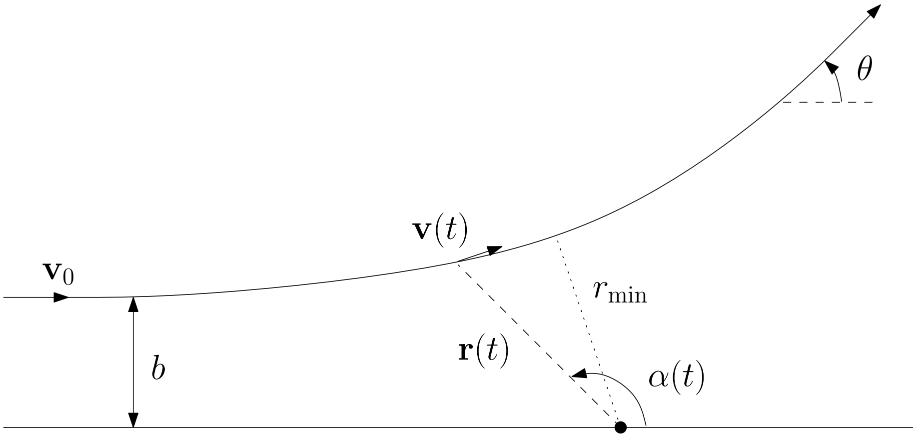
$$\sigma = \int \frac{d\sigma}{d\Omega} d\Omega = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \frac{d\sigma}{d\Omega} \sin\theta d\theta d\phi. \tag{6.2}$$

The dimension of σ is area as well. It corresponds to the total area the incident particles are passing through that will cause scattering. Put differently: imagine an area of size σ in the incident flux of the particles. The number of particles passing through this area will be equally large as the number of particles that ultimately are scattered in some direction. For instance, for scattering between hard spheres of radius R, we have $\sigma = 4\pi R^2$. For point particles scattering on a sphere of radius R, we would obtain $\sigma = \pi R^2$. We will primarily stick to central potentials $V(\boldsymbol{r}) = V(|\boldsymbol{r}|)$ which thus do not depend on the azimuthal angle ϕ due to symmetry.

B. Briefly about the classical scattering cross section

We will look at the classical case prior to the QM treatment. For a central potential (spherically symmetric), the trajectory of the particle will lie in a plane and is characterized by two quantities:

- The velocity v_0 far away from the scattering center.
- The impact parameter b (see figure below).



The scattering angle θ is determined by v_0 and b, as we show below. Assuming that $V(r \to \infty) = 0$, energy conservation gives

$$E = \frac{1}{2}mv_0^2 = \frac{1}{2}mv(t)^2 + V[(r(t)] = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\alpha}^2) + V(r), \tag{6.3}$$

and conservation of angular momentum (due to rotational symemtry) gives $L = mbv_0 = m|\boldsymbol{r} \times \boldsymbol{v}| = mr^2\dot{\alpha}$ where α characterizes the angle of the instantaoues point along the trajectory. Now, express E via L:

$$E = \frac{1}{2}m\dot{r}^2 + \frac{L^2}{2mr^2} + V(r), \tag{6.4}$$

and use that $\dot{r} = \frac{dr}{dt} = \frac{dr}{d\alpha}\frac{L}{mr^2}$. Combine these two equations to obtain:

$$d\alpha = \pm\frac{L/r^2}{\sqrt{2mE - 2mV(r) - L^2/r^2}}dr. \tag{6.5}$$

We integrate this expression from $r = r_{\text{min}}$ to $r = \infty$. Since r_{min} by definition is given by $dr/d\alpha = 0$, the $\sqrt{\ldots}$ must be zero there. The ± sign indicates whether $d\alpha/dr$ is positive or negative, which depends on the nature of the potential.

For repulsive forces (as shown in our previous figure), the change in α when going from r_{min} to $r = \infty$ is then $(\pi - \theta)/2$, while for attractive forces it would be $(\pi + \theta)/2$. The integration thus provides:

$$\frac{1}{2}(\pi \pm \theta) = \int_{r_{\text{min}}}^{\infty} \frac{L/r^2}{\sqrt{2mE - 2mV(r) - L^2/r^2}}dr. \tag{6.6}$$

Using $L = mbv_0$ and $E = \frac{1}{2}mv_0^2$, we get:

$$\frac{1}{2}(\pi \pm \theta) = \int_{r_{\text{min}}}^{\infty} \frac{b/r^2}{\sqrt{1 - V(r)E^{-1} - b^2r^{-2}}}dr. \tag{6.7}$$

This relation defines the connection between b and θ: knowing b of the incident particle, we can compute θ. A certain interval db corresponds to an interval $d\theta$ according to:

$$db = \left|\frac{db(\theta)}{d\theta}\right|d\theta = \left|\frac{db(\theta)}{d\theta}\right|\frac{d\Omega}{2\pi\sin\theta}. \tag{6.8}$$

Absolute values have been introduced since $db/d\theta$ is often negative: a large impact parameter causes a smaller scattering angle. To compute the cross section, we note that there is an area $2\pi b\ db$ contained between the impact parameters b and $b + db$. The number of particles passing through this area per unit time is $j_{\text{in}} 2\pi b\ db$. According to our above treatment, the number of particles scattered into $d\Omega$ per unit time is then:

$$j_{\text{in}} 2\pi b\ db = j_{\text{in}} b \Big|\frac{db}{d\theta}\Big| \frac{d\Omega}{\sin\theta}. \tag{6.9}$$

Using our definition of the differential scattering cross section, we obtain the final result:

$$\frac{d\sigma}{d\Omega} = \frac{b(\theta)}{\sin\theta}\Big|\frac{db(\theta)}{d\theta}\Big|. \tag{6.10}$$

To compute $d\sigma/d\Omega$, one thus has to identify $b(\theta)$.

Example 10. Coulomb-potential scattering. Two charges Ze and $Z'e$ interact via the well-known $V(r) = \frac{ZZ'e^2}{4\pi\epsilon_0 r}$. To find $b(\theta)$, we then have to compute:

$$\frac{1}{2}(\pi \pm \theta) = \int_{r_{\text{min}}}^{\infty} \frac{b/r^2}{\sqrt{1 - ZZ'e^2/(4\pi\epsilon_0 E r) - b^2 r^{-2}}} dr. \tag{6.11}$$

Upper sign: attraction ($ZZ' < 0$). Lower sign: repulsion ($ZZ' > 0$). Introducing $x = b/r$ and $g = ZZ'e^2/(8\pi\epsilon_0 E b)$, we obtain

$$\frac{1}{2}(\pi \pm \theta) = \int_0^{\sqrt{1+g^2}-g} \frac{dx}{\sqrt{1 + g^2 - (x+g)^2}}. \tag{6.12}$$

As commented on previously, x_{max} (or equivalently r_{min}) is determined by $\sqrt{\ldots} = 0$. This integral can be evaluated and yields (after rearranging the equation):

$$b = \mp \frac{ZZ'e^2}{8\pi\epsilon_0 E} \cot(\theta/2). \tag{6.13}$$

Thus, after differentiating b with respect to θ, we obtain the differential scattering cross section (known as the Rutherford cross section):

$$\frac{d\sigma}{d\Omega} = \Big(\frac{ZZ'e^2}{16\pi\epsilon_0 E}\Big)^2 \frac{1}{\sin^4(\theta/2)}. \tag{6.14}$$

For small angles θ, $d\sigma/d\Omega \propto 1/\theta^4$, causing the integral to diverge

$$\sigma = \int_0^{\pi} \frac{d\sigma}{d\Omega} 2\pi \sin\theta d\theta \to \infty. \tag{6.15}$$

Small θ corresponds to large impact parameter b, and so this result reflects the fact that the Coulomb-potential has infinite range, causing scattering of *all* incident particles.

More generally, a potential $V(r) \neq 0$ for $r \leq a$ and $V(r) = 0$ for $r > a$ will have $\sigma = \pi a^2$ classically. In QM, this is different: σ can be infinite or finite when the range $a \to \infty$, depending on how fast the potential $V(r)$ goes to zero when $r \to \infty$. The Coulomb-interaction is arguably the most important interaction in physics, and we will treat it quantum mechanically in what follows.

C. Scattering as a stationary problem

We have seen an example of such a scenario (scattering as a stationary problem) in elementary QM courses: scattering on a potential barrier in 1D. We thus seek the solution of the time independent SE:

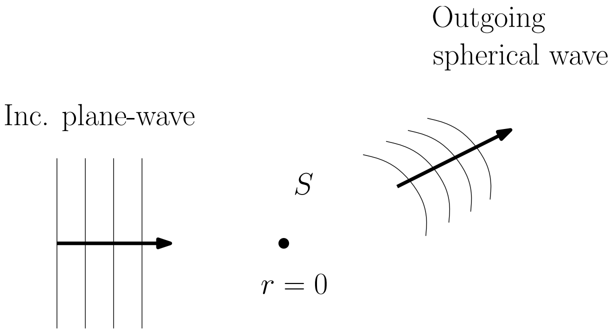
$$\Big[\frac{-\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r})\Big]\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}), \tag{6.16}$$

and use appropriate boundary conditions for the solution ψ in order to describe an incident flux of particles and an outgoing stream of scattered particles. Let $E = \hbar^2 k^2/2m$ and $U(\boldsymbol{r}) \equiv 2mV(\boldsymbol{r})/\hbar$, and obtain

$$(\nabla^2 + k^2)\psi = U\psi. \tag{6.17}$$

Near the scattering center S, the behavior of ψ may be complicated. However, for $r \to \infty$ we can neglect $U(\boldsymbol{r})$ and the resulting free particle solution should then describe an incident plane-wave and radially outgoing particles:

$$\psi \simeq \psi_{\text{in}} + \psi_{\text{scatt}} \text{ for } r \to \infty. \tag{6.18}$$



The incident wave: $\psi_{\text{in}} = C\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$ where C is a constant and $\hbar k = p_i = \sqrt{2mE}$.

The scattered wave: Must be a spherical wave with the same energy (wavenumber k) as the incident one, hence $\psi_{\text{scatt}} = Cf(\theta,\phi)\frac{\mathrm{e}^{\mathrm{i}kr}}{r}$.

FOSS

Sharp Minds - Bright Ideas!

Employees at FOSS Analytical A/S are living proof of the company value - First - using new inventions to make dedicated solutions for our customers. With sharp minds and cross functional teamwork, we constantly strive to develop new unique products - Would you like to join our team?

FOSS works diligently with innovation and development as basis for its growth. It is reflected in the fact that more than 200 of the 1200 employees in FOSS work with Research & Development in Scandinavia and USA. Engineers at FOSS work in production, development and marketing, within a wide range of different fields, i.e. Chemistry, Electronics, Mechanics, Software, Optics, Microbiology, Chemometrics.

The Family owned FOSS group is the world leader as supplier of dedicated, high-tech analytical solutions which measure and control the quality and production of agricultural, food, pharmaceutical and chemical products. Main activities are initiated from Denmark, Sweden and USA with headquarters domiciled in Hillerød, DK. The products are marketed globally by 23 sales companies and an extensive net of distributors. In line with the corevalue to be 'First', the company intends to expand its market position.

We offer
A challenging job in an international and innovative company that is leading in its field. You will get the opportunity to work with the most advanced technology together with highly skilled colleagues.

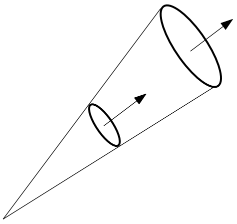
Read more about FOSS at www.foss.dk - or go directly to our student site www.foss.dk/sharpminds where you can learn more about your possibilities of working together with us on projects, your thesis etc.

Dedicated Analytical Solutions

FOSS
Slangerupgade 69
3400 Hillerød
Tel. +45 70103370

www.foss.dk

The factor $1/r$ ensures that the outgoing current density j_{scatt} is proportional to $1/r^2$. Now, since the surface element corresponding to a solid angle element $d\Omega$ increases with distance as $r^2 d\Omega$, this means that for large distances it is the same number of particles passing through any cross-section of the given, solid angle element, as expected and shown in the figure.



The factor $f(\theta, \phi)$ determines the angular distribution and is known as the scattering amplitude, which in turn is determined by $V(\boldsymbol{r})$. We shall return to this issue. Let us now determine the differential scattering cross section expressed in terms of $f(\theta, \phi)$. We know that a quantum mechanical probability current density is given as:

$$\boldsymbol{j} = \text{Re}\{\psi^* \frac{\hbar}{\text{i}m} \nabla \psi\}. \tag{6.19}$$

This means that

$$\begin{aligned} j_{\text{in}} &= \text{Re}\{\psi^*_{\text{in}} \frac{\hbar}{\text{i}m} \nabla \psi_{\text{in}}\} = \frac{\hbar k}{m} |C|^2, \\ j_{\text{scatt}} &= |C|^2 |f(\theta, \phi)|^2 \frac{\hbar k}{m r^2}. \end{aligned} \tag{6.20}$$

Since the definition of $d\sigma$ is:

$$d\sigma = \frac{j_{\text{scatt}} r^2 d\Omega}{j_{\text{in}}}, \tag{6.21}$$

we obtain by insertion:

$$\boxed{\frac{d\sigma}{d\Omega} = |f(\theta, \phi)|^2.}$$

Since C turned out to be insignificant, we set $C = 1$ in what follows, for simplicity. Note how we have obtained an expression for $d\sigma/d\Omega$ using only the asymptotic (large r) form of the wavefunction. Summarizing the idea so far:

- We seek a solution of the time independent SE

$$(\nabla^2 + k^2)\psi(\boldsymbol{r}) = U(\boldsymbol{r})\psi(\boldsymbol{r}), \tag{6.22}$$

- For large r, the solution should have the form

$$\psi(\boldsymbol{r}) \simeq \text{e}^{\text{i}\boldsymbol{k}\cdot\boldsymbol{r}} + f(\theta, \phi) \frac{\text{e}^{\text{i}kr}}{r}. \tag{6.23}$$

- The diff. scattering cross section is then:

$$\frac{d\sigma}{d\Omega} = |f(\theta, \phi)|^2. \tag{6.24}$$

The remaining task is to determine f. First, a few comments:

1. We have here assumed elastic scattering, thus neglecting the possibility of the particles making energy transitions during the collision.
2. We have assumed free particle behavior at $r \to \infty$. For potentials with infinite range, it is essential how fast $V \to 0$ when $r \to \infty$. It turns out that if $rV(r) \to 0$ for $r \to \infty$, we obtain the free particle asymptotic behavior. The Coulomb-potential *does not* satisfy this and we shall later see how this influences the asymptotic form.

3. In a real experiment, one does not send in an infinite plane-wave toward S, but rather a beam that is collimated (focused) in space. The localized nature (let l_i be the width of the beam in the i-direction) causes an uncertainty in the momentum $\Delta p_i \simeq \hbar/l_i$. However, l_i will usually be much larger than atomic distances, and the lack of precision in momentum should thus be negligible compared to the change in momentum (direction) caused by the potential. We may thus disregard finite-size effect of the beam and model it with a plane-wave.

4. In practice, one scatters particles on a macroscopic collection of particles rather than a single scattering center S, *e.g.* a gas of particles. To use our approach, the thickness of the target has to be large enough to cause sufficient scattering intensity, but small enough to keep multiple scattering at a minimum.

D. Integral equation for the scattering amplitude

Our strategy here will be to transform the SE into an integral equation in order to incorporate the correct asymptotic behavior $\psi(\boldsymbol{r}) \simeq \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} + f(\theta,\phi)\mathrm{e}^{\mathrm{i}kr}/r$. This transformation is done with the aid of the Green function $G(\boldsymbol{r}-\boldsymbol{r}')$, defined by

$$(\nabla^2 + k^2)G(\boldsymbol{r}-\boldsymbol{r}') = \delta(\boldsymbol{r}-\boldsymbol{r}'). \tag{6.25}$$

If G is known, then the SE is equivalent to:

$$\psi(\boldsymbol{r}) = \psi_0(\boldsymbol{r}) + \int G(\boldsymbol{r}-\boldsymbol{r}')U(\boldsymbol{r}')\psi(\boldsymbol{r}')d\boldsymbol{r}', \tag{6.26}$$

where ψ_0 is a general solution of the homogeneous equation $(\nabla^2 + k^2)\psi_0 = 0$. To establish this equivalence, operate with $\nabla^2 + k^2$ on Eq. (6.26):

$$(\nabla^2 + k^2)\psi = 0 + \int \delta(\boldsymbol{r}-\boldsymbol{r}')U(\boldsymbol{r}')\psi(\boldsymbol{r}')d\boldsymbol{r}' = U(\boldsymbol{r})\psi(\boldsymbol{r}) \tag{6.27}$$

which is precisely the SE. The second order differential equation for G has two independent solutions:

$$G(\boldsymbol{r}-\boldsymbol{r}') = -\frac{\mathrm{e}^{\pm\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{4\pi|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{6.28}$$

To see this, it is sufficient to demonstrate that $G(\boldsymbol{r}) = -\frac{\mathrm{e}^{\pm\mathrm{i}kr}}{4\pi r}$ satisfies $(\nabla^2 + k^2)G(\boldsymbol{r}) = \delta(\boldsymbol{r})$. Since $\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} +$ angular derivatives, we obtain for $r > 0$:

$$\frac{d}{dr}\frac{\mathrm{e}^{\pm\mathrm{i}kr}}{r} = \Big(\frac{\pm\mathrm{i}k}{r} - \frac{1}{r^2}\Big)\mathrm{e}^{\pm\mathrm{i}kr} \tag{6.29}$$

and

$$\frac{d^2}{dr}\frac{\mathrm{e}^{\pm\mathrm{i}kr}}{r} = \Big(-\frac{k^2}{r} \mp \frac{2\mathrm{i}k}{r^2} + \frac{2}{r^3}\Big)\mathrm{e}^{\pm\mathrm{i}kr}. \tag{6.30}$$

Combined, this yields $(\nabla^2 + k^2)\frac{\mathrm{e}^{\pm\mathrm{i}kr}}{r} = 0$. This is consistent since $\delta(\boldsymbol{r}) = 0$ for $\boldsymbol{r} \neq 0$. To justify the presence of the δ-function, we integrate $(\nabla^2 + k^2)G$ over a spherical volume with radius R by using the formula:

$$\int_V \nabla^2 G d\boldsymbol{r} = \int_{f(V)} \nabla G \cdot d\boldsymbol{f} = 4\pi R^2(\partial_r G)_{r=R} \tag{6.31}$$

where $f(V)$ is the surface of the volume V. We then get:

$$\begin{aligned}\int_{r\leq R}(\nabla^2 + k^2)G &= 4\pi R^2\Big(\frac{\pm\, k\mathrm{e}^{\pm\mathrm{i}kr}}{-4\pi R} + \frac{\mathrm{e}^{\pm\mathrm{i}kr}}{4\pi R^2}\Big) + k^2\int_0^R \frac{\mathrm{e}^{\pm\mathrm{i}kr}}{-4\pi r}4\pi r^2 dr \\ &= \mp\mathrm{i}kR\mathrm{e}^{\pm\mathrm{i}kr} + \mathrm{e}^{\pm\mathrm{i}kr} + k^2\Big[\frac{\mp r}{\mathrm{i}k}\mathrm{e}^{\pm\mathrm{i}kr} - \frac{\mathrm{e}^{\pm\mathrm{i}kr}}{k^2}\Big]_0^R = 1.\end{aligned} \tag{6.32}$$

Since the integral over $(\nabla^2 + k^2)G$ is 1 for *any* finite radius R, we must have $(\nabla^2 + k^2)G = \delta(\boldsymbol{r})$, which completes the proof. Now that we know exactly what $G(\boldsymbol{r}-\boldsymbol{r}')$ is, we can insert it into Eq. (6.26) in order to find $\psi(\boldsymbol{r})$. The choice of $\psi_0(\boldsymbol{r})$ is dictated by the boundary conditions and the fact that it has to satisfy

$(\nabla^2 + k^2)\psi_0 = 0$. We therefore set $\psi_0(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$ since $U = 0$ should give precisely this wavefunction as there is no scattering in this case.

Secondly, we choose the $+$ solution for G so that we recover the correct form $\mathrm{e}^{\mathrm{i}kr}/r$ for large r. With these two choices, the solution for $\psi(\boldsymbol{r})$ then becomes:

$$\psi(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{1}{4\pi}\int \frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} U(\boldsymbol{r}')\psi(\boldsymbol{r}')d\boldsymbol{r}'. \tag{6.33}$$

Note that $|\boldsymbol{r} - \boldsymbol{r}'| \simeq r$ for large r. If we now look at the behavior of the above equation in the large-r limit, we will be able to identify $f(\theta, \phi)$ by comparing directly with the form $\psi(\boldsymbol{r}) \simeq \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} + \mathrm{e}^{\mathrm{i}kr} f(\theta, \phi)/r$.

First, we do the large-r expansion more accurately. We have:

$$k|\boldsymbol{r}-\boldsymbol{r}'| = k\sqrt{r^2 - 2\boldsymbol{r}\cdot\boldsymbol{r}' + (r')^2} = kr\sqrt{1 - \frac{2\boldsymbol{r}\cdot\boldsymbol{r}'}{r^2} + \frac{(r')^2}{r^2}} = kr - \boldsymbol{k}'\cdot\boldsymbol{r}' + \mathcal{O}(1/r), \tag{6.34}$$

where $\boldsymbol{k}' \equiv k\boldsymbol{r}/r$ points in the direction that the particle has after scattering. The momentum of the final state is thus $\boldsymbol{p}_f = \hbar\boldsymbol{k}'$. Note that $|\boldsymbol{p}_f| = \hbar|\boldsymbol{k}| = \hbar|\boldsymbol{k}'|$: conservation of momentum. Using our expansion, the integral equation then takes the form

$$\psi(\boldsymbol{r}) \simeq \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{1}{4\pi}\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\int \mathrm{e}^{-\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}'} U(\boldsymbol{r}')\psi(\boldsymbol{r}')d\boldsymbol{r}'. \tag{6.35}$$

Now, we can finally read out the scattering amplitude:

$$\boxed{f(\theta, \phi) = -\frac{1}{4\pi}\int \mathrm{e}^{-\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}'} U(\boldsymbol{r}')\psi(\boldsymbol{r}')d\boldsymbol{r}'.}$$

It may appear as if we still have not accomplished much since this expression still depends on the unknown $\psi(\boldsymbol{r})$. However, it turns out that our formulation is still useful, because we have now set up the problem in a manner which makes it suitable for an *iterative treatment*.

E. Born-approximation

If the scattering potential is weak (and we shall later specify what this means quantitatively), we can solve our integral equation for ψ by iteration. The n-th approximation is obtained by using the $(n-1)$-th approximation on the r.h.s. of

$$\psi(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{1}{4\pi}\int \frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}U(\boldsymbol{r}')\psi(\boldsymbol{r}')d\boldsymbol{r}'. \tag{6.36}$$

The most basic approximation, $\psi^{(0)}(\boldsymbol{r})$, is simply to set it equal to the incident plane wave. Thus, we obtain

$$\begin{aligned}
\psi^{(0)}(\boldsymbol{r}) &= \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}},\\
\psi^{(1)}(\boldsymbol{r}) &= \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{1}{4\pi}\int \frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}U(\boldsymbol{r}')\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}'}d\boldsymbol{r}',\\
\psi^{(2)}(\boldsymbol{r}) &= \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{1}{4\pi}\int \frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}U(\boldsymbol{r}')\psi^{(1)}(\boldsymbol{r}')d\boldsymbol{r}',\\
\psi^{(3)} &= \ldots
\end{aligned} \tag{6.37}$$

and so forth. In this manner, we can obtain better and better approximations for $f(\theta,\phi)$ by inserting approximations for $\psi(\boldsymbol{r})$. This expansion is known as the Born-approximation and one often settles for the lowest order correction. We now examine this in more detail.

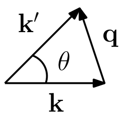
First order Born-approximation.
Using $\psi^{(0)} = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$, we obtain

$$f^B(\theta,\phi) = -\frac{1}{4\pi}\int \mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}U(\boldsymbol{r})d\boldsymbol{r}, \tag{6.38}$$

where the B superscript indicates that this result has been obtained in the first-order Born-approximation. Introducing $\boldsymbol{q} = \boldsymbol{k}' - \boldsymbol{k}$ and reinstating $U = 2mV/\hbar^2$, we get:

$$\boxed{f^B = -\frac{m}{2\pi\hbar^2}\int V(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}d\boldsymbol{r}.}$$

In other words, the scattering amplitude f^B is essentially the Fourier-transform of the potential. The physical meaning of $\boldsymbol{q}$ is that it is the momentum-transfer during the collision: $q = 2k\sin(\theta/2)$ according the figure.



Note that so far, we have not made any assumption about the potential being spherically symmetric. If it is, however, we may simplify the expression for f^B as follows. Let $V(\boldsymbol{r}) = V(r)$ and let $\boldsymbol{r}$ point along the polar axis. We then obtain

$$\int_{\nu=0}^{\pi}\int_{\zeta=0}^{2\pi} \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}d\zeta\sin\nu d\nu = 2\pi\Big[\frac{\mathrm{e}^{-\mathrm{i}qr\cos\nu}}{\mathrm{i}qr}\Big]_{\nu=0}^{\nu=\pi} = \frac{4\pi\sin(qr)}{qr}. \tag{6.39}$$

The result for f^B is then:

$$f^B(\theta) = -\frac{2m}{\hbar^2 q}\int_0^\infty V(r)\sin(qr)rdr. \tag{6.40}$$

Note that ζ and ν are just integration variables without any special significance. In the forward scattering case $(\theta = 0)$, f^B becomes independent on q and thus the energy of the particle. This might appear strange at first glance and in fact it is physically incorect: it is an artifact of the perturbation expansion of the Born treatment.

Going to 2nd order in the perturbation $V(r)$ fixes the problem. The point is that one must account for interference between the incoming wave and the outgoing wave for $\theta = 0$ in order to correctly describe forward scattering.

In contrast, this interference is not as important for $\theta \neq 0$. For such directions, the oscillating term $\sin(qr)$ renders the integral small when $qa \gg 1$ where a is a measure for the spatial range of the potential. This means that for high energies (large k), the differential scattering cross section is very small except when $qa = 2ka\sin(\theta/2) \simeq 1$, which for large k means $\theta \simeq 1/(ka)$. We then conclude that high-energy particles do not change their direction much, keeping their trajectory close to $\theta \simeq 0$.

For the total scattering cross section, we have that $\sigma^B = \int \frac{d\sigma^B}{d\Omega} d\Omega = \int |f^B(\theta,\phi)|^2 d\Omega$. For a spherically symmetric potential, we obtain

$$d\Omega = 2\pi \sin\theta d\theta = 4\pi \sin(\theta/2)\cos(\theta/2) d\theta = \frac{2\pi}{k^2} q\, dq \tag{6.41}$$

where we utilized that $2k\sin(\theta/2) = q$ so that $dq = k\cos(\theta/2)d\theta$. σ^b can then be obtained by integrating over q:

$$\sigma^B = \frac{2\pi}{k} \int_0^{2k} |f^B(q)|^2 q\, dq. \tag{6.42}$$

Since the integral grows as $E = \hbar^2k^2/2m$ increases, σ^B cannot decrease faster than $1/E$. More precisely, if the integral converges at high energies, one obtains $\sigma^B \propto 1/E$. We are treating this problem non-relativistically, so "high energies" still means that $E \ll mc^2$.

When is the Born-approximation valid?

The iteration procedure that we have utilized is based on the assumption that the incident plane wave is not severly altered. In effect, we require that $|\psi(\boldsymbol{r}) - \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}| \ll 1$. Using our expression for ψ in the Born-approximation provides:

$$\left| \frac{1}{4\pi} \int \frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} U(\boldsymbol{r}') \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}'} d\boldsymbol{r}' \right| \ll 1. \tag{6.43}$$

The modification of the incident wave is expected to be largest at the scattering center $\boldsymbol{r} = 0$, so the strictest requirement is:

$$\left| \frac{1}{4\pi} \int \frac{\mathrm{e}^{\mathrm{i}kr'}}{r'} U(\boldsymbol{r}') \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}'} d\boldsymbol{r}' \right| \ll 1. \tag{6.44}$$

If we want an *even stricter* requirement, we take the absolute value of all factors in the integrand:

$$\int_0^\infty |U(r)| r\, dr \ll 1 \tag{6.45}$$

where we used that $d\boldsymbol{r} = 4\pi r^2 dr$ for $U(\boldsymbol{r}) = U(r)$. When this inequality is satisfied, the Born-approximation is expected to be good for all energies.

Example 11. Bound-state in a constant potential. For a constant potential V_0 with range R, the criterion of validity takes the form

$$\frac{m|V_0|R^2}{\hbar^2} \ll 1. \tag{6.46}$$

At the same time, we know that a negative potential $-|V_0|$ can bind states if $m|V_0|R^2 > \pi^2\hbar^2/8$ from introductory QM. This is in agreement with the criterion: we expect the Born-approximation to be valid when $|V_0|$ is sufficiently weak to be unable to bind a particle with mass equal to the incident particles.

For a finite-range potential with finite magnitude, one can *always use* the Born-approximation at *sufficiently high energies*. To see this, let us go back to the k-dependent criterion

$$\left| \frac{1}{4\pi} \int \frac{\mathrm{e}^{\mathrm{i}kr'}}{r'} U(\boldsymbol{r}') \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}'} d\boldsymbol{r}' \right| \ll 1. \tag{6.47}$$

Assume now a spherically symmetric potential $U(\boldsymbol{r}) = U(r)$, in which case we can perform the angular integration:

$$\int_0^\pi \mathrm{e}^{\mathrm{i}kr'\cos\nu}\sin\nu\, d\nu = \frac{2\sin(kr')}{kr'}. \tag{6.48}$$

This then gives:

$$\left|\frac{1}{k}\int_0^\infty U(\boldsymbol{r}')\mathrm{e}^{\mathrm{i}kr'}\sin(kr')dr'\right| \ll 1. \tag{6.49}$$

Since $|\mathrm{e}^{\mathrm{i}kr'}\sin(kr')| \leq 1$, the criterion

$$k \gg \int_0^\infty |U(r)|dr \tag{6.50}$$

is sufficient to guarantee that the original k-dependent criterion is fulfilled. We conclude that for large enough k, we can always fulfill the above equation. Note that using the Born-approximation, we have actually obtained the same result for $d\sigma/d\Omega$ for a weak potential V as we did using time dependent perturbation theory in chapter 4:

$$\frac{d\sigma}{d\Omega} = \left|\frac{m}{2\pi\hbar}\int V(\boldsymbol{r})\mathrm{e}^{\mathrm{i}(\boldsymbol{p}_i - \boldsymbol{p}_f)\cdot\boldsymbol{r}/\hbar}d\boldsymbol{r}\right|^2, \tag{6.51}$$

which is reasonable since the Born-approximation is good when the potential is weak or the particle energy is high.

Example 12. Scattering on the Yukawa-potential. Let us apply our scattering framework on a screened Coulomb-potential which is known as a Yukawa potential:

$$V(r) = \frac{ZZ'e^2}{4\pi\epsilon_0 r}\mathrm{e}^{-\alpha r}. \tag{6.52}$$

Here, α determines the screening radius. Using our derived result for the Born scattering amplitude gives:

$$f^B(\theta) = -\frac{2m}{\hbar^2 q}\frac{ZZ'e^2}{4\pi\epsilon_0}\int_0^\infty \mathrm{e}^{-\alpha r}\sin(qr)dr = -\frac{2m}{\hbar}^2\frac{ZZ'e^2}{4\pi\epsilon_0}\frac{1}{\alpha^2+q^2}. \tag{6.53}$$

Using the relations introduced previously: $q = 2k\sin(\theta/2),\ \ k = p/\hbar = \sqrt{2mE}/\hbar$, we can write down the differential scattering cross section:

$$\frac{d\sigma^B}{d\Omega} = |f^B|^2 = \Big(\frac{ZZ'e^2/4\pi\epsilon_0}{\alpha^2\hbar^2(2m)^{-1} + 4E\sin^2(\theta/2)}\Big)^2. \tag{6.54}$$

It is interesting to note that since $d_\Omega\sigma^B$ is finite for all angles when $\alpha \neq 0$, the total σ^B will also be finite. This is *in contrast* to the classical value σ for this potential which becomes infinite. In the limit $\alpha \to 0$, we obtain the usual unscreened Coulomb-potential. Remarkably, the QM Born-result for $d_\Omega\sigma \equiv d\sigma/d\Omega$ is not only identical to the classical result, but it is even identical to the *exact* QM Coulomb cross section! The derivation is not shown here, but one finds in the exact treatment that

$$f^{\text{exact}}(\theta) = \frac{n}{2k\sin^2(\theta/2)}\mathrm{e}^{-2\mathrm{i}\ln\sin(\theta/2)+\mathrm{i}\delta}, \tag{6.55}$$

where we defined

$$n = \frac{m}{k\hbar^2}\frac{ZZ'e^2}{4\pi\epsilon_0} \tag{6.56}$$

and where δ is a constant (independent on θ). Since $|\mathrm{e}^{\mathrm{i}\delta}| = 1$, this phase-factor has no consequence for the cross section. However, we will later show that when scattering *identical* particles on each other, it will have an effect. The result that

$$\frac{d\sigma^{\text{classical}}}{d\Omega} = \frac{d\sigma^B}{d\Omega} = \frac{d\sigma^{\text{exact}}}{d\Omega} \tag{6.57}$$

for the Coulomb-potential must be regarded as a coincidence, since the criteria we listed for the Born-approximation are not expected to be valid for the Coulomb-potential.

Elastic scattering on atoms.
If the electrons are fast (energetic), we can treat scattering on neutral atoms via the potential

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r} + \frac{e^2}{4\pi\epsilon_0}\int \frac{n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}d\boldsymbol{r}'. \tag{6.58}$$

This consists of the Coulomb repulsion from the core $+Ze$ in addition to the potential from the electron distribution $-en(\boldsymbol{r})$. Charge neutrality dictates that $\int n(\boldsymbol{r})d\boldsymbol{r} = Z$. The reason for why the electrons must be fast in order for us to use the above potential is that the true antisymmetrized ψ gives a correction to the result otherwise. ψ must be antisymmetrized for scattering of electrons on an electron, since these are identical particles quantum mechanically. Recall that r is the relative coordinate between the potential and the scattered particles.

We use the Born-approximation, meaning that the incident electron E satisfies $13.6Z^2$ eV $\ll E \ll$ 500 000 eV: it is much larger than the typical potential energy scale, while still non-relativistic ($m_e c^2 \simeq 0.5$ MeV). Now, we seek the scattering amplitude $f^B = -\frac{m}{2\pi\hbar^2}\int V(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}d\boldsymbol{r}$. Introducing $\boldsymbol{s} = \boldsymbol{r} - \boldsymbol{r}'$ and using our result for the Yukawa-potential without screening $(\alpha = 0)$, we get:

$$\int \frac{\mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}}{|\boldsymbol{r}-\boldsymbol{r}'|}d\boldsymbol{r} = \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}'}\frac{4\pi}{q^2}. \tag{6.59}$$

It follows that

$$f^B = \frac{e^2}{4\pi\epsilon_0}\frac{2m}{\hbar^2}\frac{Z - F(\boldsymbol{q})}{q^2}, \tag{6.60}$$

where $F(\boldsymbol{q}) = \int n(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}d\boldsymbol{r}$ is known as the atom form-factor. It is the Fourier-transformation of the electron distribution. If $n(\boldsymbol{r})$ is spherically symmetry, it follows that $F(\boldsymbol{q}) = F(q)$. With $q = 2k\sin(\theta/2)$ and $E = \hbar^2k^2/2m$ as usual, the differential scattering cross section becomes:

$$\frac{d\sigma^B}{d\Omega} = |f|^2 = \Big(\frac{e^2}{16\pi\epsilon_0 E}\Big)^2 \sin^{-4}(\theta/2)[Z - F(2k\sin(\theta/2))]^2. \tag{6.61}$$

By measuring $d\sigma^B/d\Omega$, we can thus obtain information about $F(q)$ and, in turn, the electronic distribution of the atom. The idea is thus to scatter a simple particle on a complex structure (potential) to gain info about the complex structure.

Two particular limiting cases are of interest:

1. When the scattering angle θ *is not small*, $q = 2k\sin(\theta/2)$ is sizable when k is large (energetic electrons). The result is that F becomes small since the integrand oscillates around zero. Quantitatively, this requires that $1/q \ll$ atomic dimensions, i.e. ~ 1 Å. If we thus can neglect F compared to Z, $d_\Omega\sigma$ is essentially the Rutherford cross section. This result makes sense physically: a particle with high E can only scatter a large angle θ if it comes close to the core.

2. In the opposite regime, *for very small angles*, we can expand F in powers of q:

$$F(q) = \int n(r)[1 - \mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r} - \frac{1}{2}(\boldsymbol{q}\cdot\boldsymbol{r})^2 + \ldots]d\boldsymbol{r} = Z - \frac{q^2}{6}\int r^2 n(r)d\boldsymbol{r}. \tag{6.62}$$

   The second term $\propto \mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}$ vanishes due to symmetery. For the third term, we used that the integral with $q_x^2x^2 + q_y^2y^2 + q_z^2z^2$ is 1/3 of the integral with $(q_x^2 + q_y^2 + q_z^2)r^2 = q^2r^2$. Define now the average atomic radius R:

$$R^2 = \frac{\int r^2 n(r)d\boldsymbol{r}}{\int n(r)d\boldsymbol{r}} = \frac{1}{Z}\int r^2 n(r)d\boldsymbol{r}. \tag{6.63}$$

   This yields $Z - F(q) \simeq ZR^2q^2/6$, which in turn is $\ll Z$ for small angles (small q). This means that

$$\frac{d\sigma^B}{d\Omega} \simeq \Big(\frac{ZR^2}{3a_0}\Big)^2 \tag{6.64}$$

   with $a_0 = 4\pi\epsilon_0\hbar^2/me^2$. The scattering is then independent on θ and small when θ is small. This may be physically interpreted as the electron cloud effectively screening the core Ze at small scattering angles (classically, this corresponds to a large impact parameter).

F. The method of partial waves

So far, we have seen that the Born-approximation is good when E of the incident particle is large. Now, we will consider a method which is good in the *opposite case*, namely the partial wave method which is useful for low energies (for instance scattering of sound) and was developed in 1927 by Holtsmark and Faxen.

Scattering amplitude
We know that the energy eigenfunctions for a spherically symmetric potential $V(r)$ can be written generally as

$$\psi(r,\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{lm}R_l(r)Y_{lm}(\theta,\phi) \tag{6.65}$$

where $k^2 = 2mE/\hbar^2$ and $U(r) = 2m\hbar^{-2}V(r)$ where R satisfies:

$$\frac{d^2}{dr^2}(rR_l) + \Big[k^2 - U(r) - \frac{l(l+1)}{r^2}\Big](rR_l) = 0. \tag{6.66}$$

We are interested in $E > 0$ (continuous part of the spectrum) and cylinder symmetric (no ϕ-dependence) solutions, as is often the case in scattering problems. Choose $\hat{\boldsymbol{z}}$ as the incident axis. Eq. (6.65) is independent on ϕ when $m = 0$. The function $Y_{lm}(\theta, \phi)$ then reduces to Legendre-polynomials $P_l(\cos\theta)$. These form a complete set for cylinder symmetric functions, so that:

$$\psi(r,\theta) = \sum_{l=0}^{\infty} c_l R_l(r) P_l(\cos\theta) \tag{6.67}$$

where c_l are constants. We may then also express $f(\theta) = \sum_{l=0}^{\infty} f_l P_l(\cos\theta)$ where f_l are constants. This is the announced expansion in "partial waves", each characterized by a quantum number l. Keep in mind that $P_l(\cos\theta)$ is an eigenfunction for $\hat{\boldsymbol{L}}^2$ with eigenvalues $\hbar^2 l(l+1)$. We recall that the scattering amplitude $f(\theta)$ is defined from the asymptotic behavior: $\psi(\boldsymbol{r}) - \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \simeq f(\theta)\frac{\mathrm{e}^{\mathrm{i}kr}}{r}$ valid for $r \to \infty$. To expand f in partial waves, we must first expand $\psi(\boldsymbol{r})$ and $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$ in Legendre polynomials. Start with the incident wave:

$$\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} = \mathrm{e}^{\mathrm{i}kr\cos\theta} = \sum_{l=0}^{\infty} d_l(kr) P_l(\cos\theta). \tag{6.68}$$

Introduce $x \equiv \cos\theta$ and use the orthogonality of $P_l(\cos\theta)$:

$$\int_{-1}^{1} P_l(x) P_n(x) dx = \frac{2}{2l+1}\delta_{ln}. \tag{6.69}$$

Applied on Eq. (6.68), we obtain by multiplying with $P_n(x)$ on both sides and integrating:

$$\int_{-1}^{1} P_n(x)\mathrm{e}^{\mathrm{i}krx} dx = \sum_{l=0}^{\infty}\int_{-1}^{1} d_l(kr) P_l(x) P_n(x)\ dx. \tag{6.70}$$

Therefore, we see that

$$d_l(kr) = \frac{2l+1}{2}\int_{-1}^{1} \mathrm{e}^{\mathrm{i}krx} P_l(x) dx. \tag{6.71}$$

We want to see how this behaves for large r. To do so, consider general integrals of the form $I \equiv \int_{-1}^{1} \mathrm{e}^{\mathrm{i}sx} g(x) dx$ for large s. Consecutive partial integrations, where the exponential function is integrated, provides:

$$I = \Big[g(x)\frac{\mathrm{e}^{\mathrm{i}sx}}{\mathrm{i}s}\Big]_{-1}^{1} - \int_{-1}^{1} g'(x)\frac{\mathrm{e}^{\mathrm{i}sx}}{\mathrm{i}s} dx = \Big[g(x)\frac{\mathrm{e}^{\mathrm{i}sx}}{\mathrm{i}s} - g'(x)\frac{\mathrm{e}^{\mathrm{i}sx}}{(\mathrm{i}s)^2}\Big]_{-1}^{1} + \int_{-1}^{1} g''(x)\frac{\mathrm{e}^{\mathrm{i}sx}}{(\mathrm{i}s)^2} dx \tag{6.72}$$

and so forth. Since s is presumed to be large, we obtain smaller and smaller terms. The dominating term for large s is then:

$$I = \int_{-1}^{1} \mathrm{e}^{\mathrm{i}sx} g(x) dx = g(1)\frac{\mathrm{e}^{\mathrm{i}s}}{\mathrm{i}s} - g(-1)\frac{\mathrm{e}^{-\mathrm{i}s}}{\mathrm{i}s} + \mathcal{O}(s^{-2}). \tag{6.73}$$

In our case, $s = kr$ and $g(x) = P_l(x)$. Moreover, $P_l(1) = 1$ and $P_l(-1) = (-1)^l$ by definition. Hence, the asymptotic behavior of d_l is:

$$d_l \simeq \frac{2l+1}{2\mathrm{i}kr}[\mathrm{e}^{\mathrm{i}kr} - (-1)^l \mathrm{e}^{-\mathrm{i}kr}]. \tag{6.74}$$

Using $(-1)^l = \mathrm{e}^{\mathrm{i}\pi l}$ and $2\mathrm{i}\sin y = \mathrm{e}^{\mathrm{i}y} - \mathrm{e}^{-\mathrm{i}y}$, we rewrite this to $d_l(kr) \simeq (2l+1)\mathrm{i}^l \frac{\sin(kr - \frac{1}{2}l\pi)}{kr}$ for large r. We have now managed to identify how $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$ is expanded for large $\boldsymbol{r}$. It remains to find the asymptotic behavior of $\psi(\boldsymbol{r})$. We have $\psi(r,\theta) = \sum_{l=0}^{\infty} c_l R_l(r) P_l(\cos\theta)$ where the equation determining R_l in the limit $r \to \infty$ reads:

$$\frac{d^2}{dr^2}(rR_l) + k^2(rR_l) = 0. \tag{6.75}$$

Hence, we disregarded $l(l+1)/r^2$ and $U(r)$ [this is fine when $U(r)$ drops faster than $1/r$ for large r]. The solution of Eq. (6.75) is sine and cosine functions. With two arbitrary constants c_l and δ_l, we can write the general solution:

$$rR_l \simeq (2l+1)\mathrm{i}^l c_l \sin(kr - \pi l/2 + \delta_l) \tag{6.76}$$

for large r. We have written the solution in this form to look as similar as possible to the expansion of $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$ obtained previously. The quantity δ_l is the phase picked up by the wavefunction as a consequence of the scattering potential and is referred to as the l-th *scattering phase*. In the absence of any potential $U(r)$, one finds $\delta_l = 0$. To determine δ_l in the general case, one has to solve the radial equation for all r and then inspect R_l for large r. Inserting our expansions, we have thus found

$$\psi(\boldsymbol{r}) - \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \simeq P_l(\cos\theta)(2l+1)\mathrm{i}^l[c_l\sin(kr - l\pi/2 + \delta_l) - \sin(kr - l\pi/2)]/kr. \tag{6.77}$$

In order to finally identify $f(\theta)$, we should now focus on the conditon that the above expression should only contain spherical waves of the form $\mathrm{e}^{\mathrm{i}kr}/r$ according to the asymptotic expression for the wavefunction. This is accomplished by noting that:

$$[\ldots] = \frac{1}{2\mathrm{i}}(c_l\mathrm{e}^{\mathrm{i}\delta_l} - 1)\mathrm{e}^{\mathrm{i}(kr - l\pi/2)} - \frac{1}{2\mathrm{i}}(c_l\mathrm{e}^{-\mathrm{i}\delta_l} - 1)\mathrm{e}^{-\mathrm{i}(kr - l\pi/2)}. \tag{6.78}$$

It is clear that we must choose $c_l = \mathrm{e}^{\mathrm{i}\delta_l}$ to remove the $\mathrm{e}^{-\mathrm{i}kr}$ term, which leaves us with

$$\psi(\boldsymbol{r}) - \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \simeq \frac{\mathrm{e}^{\mathrm{i}kr}}{2\mathrm{i}kr}\sum_{l=0}^{\infty}(2l+1)(\mathrm{e}^{2\mathrm{i}\delta_l} - 1)P_l(\cos\theta). \tag{6.79}$$

Since $\mathrm{e}^{2\mathrm{i}\delta_l} - 1 = 2\mathrm{i}\mathrm{e}^{\mathrm{i}\delta_l}\sin\delta_l$, we can now identify $f(\theta)$:

$$\boxed{f(\theta) = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\mathrm{e}^{\mathrm{i}\delta_l}\sin\delta_l P_l(\cos\theta).}$$

where, as usual $d_\Omega\sigma = |f(\theta)|^2$. We have not solved the problem entirely yet, but we have established a connection between the solution of the radial equation (i.e. determining δ_l) and $f(\theta)$. We will look at a concrete application later where δ_l is determined and hence solving the problem.

Total scattering cross section.
Our expression for $f(\theta)$ determines σ:

$$\sigma = \int_0^\pi |f(\theta)|^2 2\pi \sin\theta d\theta = \frac{2\pi}{k^2}\sum_{ll'}(2l+1)(2l'+1)\mathrm{e}^{\mathrm{i}\delta_l - \mathrm{i}\delta_l'}\sin\delta_l \sin\delta_l' \int_{-1}^1 P_l(x)P_{l'}(x)dx. \tag{6.80}$$

Using the aforementioned orthogonality of $P_l(x)$, we obtain:

$$\sigma = \frac{4\pi}{k^2}\sum_{l=0}^\infty (2l+1)\sin^2\delta_l. \tag{6.81}$$

We may note an interesting relation between σ and the forward-scattering amplitude ($\theta = 0$). Setting $\theta = 0$ in $f(\theta) = \frac{1}{k}\sum_{l=0}^\infty (2l+1)\mathrm{e}^{\mathrm{i}\delta_l}\sin\delta_l P_l(\cos\theta)$ and using that $P_l(1) = 1$, we obtain $f(0) = \frac{1}{k}\sum_{l=0}^\infty (2l+1)\mathrm{e}^{\mathrm{i}\delta_l}\sin\delta_l$. It follows that we may write generally:

$$\boxed{\sigma = \frac{4\pi}{k}\mathrm{Im}\{f(0)\}.}$$

This relation is known as the *optical theorem.* We will later give a general proof of the theorem. The fact that $f(0)$ appears is related to that in order to cause scattering, the incident beam must be weakened. This is achieved via destructive interference between the incident beam and the outgoing forward-scattered beam, described precisely via $f(0)$.

Number of significant phases.
If particles with momentum $\hbar k$ approach a potential with range R, only particles with angular momentum $\hbar kR$ or less should be scattered from a classical perspective. Since the angular momentum size is $\hbar\sqrt{l(l+1)}$, we obtain that

$$\sqrt{l(l+1)} \leq kR. \tag{6.82}$$

For low energies $kR \ll 1$ (particle wavelength $2\pi/k \gg R$), we see that only $l = 0$ contributes. In this case, $f \simeq k^{-1}\mathrm{e}^{\mathrm{i}\delta_0}\sin\delta_0$ as only the $l = 0$ partial wave contributes and we obtain isotropic scattering since

$$\frac{d\sigma}{d\Omega} \simeq k^{-2}\sin^2\delta_0 \to \sigma = \frac{4\pi}{k^2}\sin^2\delta_0. \tag{6.83}$$

This shows why the partial wave method is so useful for low energies. The scattering amplitude has dimension length, and the low energy limit for f (which for finite-range potentials is independent on the angle) is often called the *scattering length* a:

$$\lim_{k\to 0} f = -a. \tag{6.84}$$

This results in $\sigma = 4\pi a^2$. To be more specific about what "low energies" means, note that $kR \ll 1$ gives:

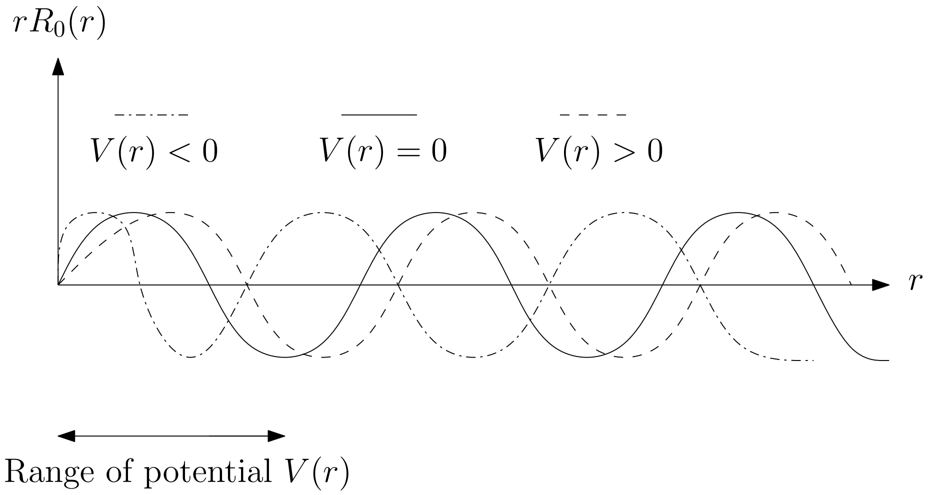
$$E = \frac{\hbar^2 k^2}{2m} \ll \frac{\hbar^2}{2mR^2} = \frac{m}{m_e}\Big(\frac{a_0}{R}\Big)^2 \times 13.6 \text{ eV} \tag{6.85}$$

by using the energy expression for the $n = 1$ Coulomb potential.

Ramsauer-Townsend effect: sign of the phase-change.
The phase-change δ_0 determines the scattering cross section at low energies. sign(δ_0) is related to the sign of the potential $U(r)$. To see this, recall that δ_0 was defined by writing the solution of

$$\frac{d}{dr^2}u_0 + [k^2 - U(r)]u_0 = 0, \tag{6.86}$$

where $u_0 = rR_0(r)$, on the form $u_0 \propto \sin(kr + \delta_0)$ for large r. If $U < 0$, the effective wavenumber $\sqrt{k^2 - U(r)}$ is larger than for $U = 0$. In turn, this means that the particle wavelength λ inside the potential becomes shorter, so that u_0 will have a stronger curvature. The wavefunction is then "pulled" closer to $r = 0$, corresponding to a positive phase-change as shown in the figure.



Conversely, a positive potential gives a negative δ_0. Outside of the range of U, the wavelength is of course the same in all cases. If the attractive potential ($U < 0$) is sufficiently strong to "pull in" the partial wave $l = 0$ to the extent that $\delta_0 = \pi$, then $\sin\delta_0 = \pi$ and $\sigma \to 0$: the scattering cross section vanishes. For a given potential, this effect (Raumsauer-Townsend) requires a specific energy. It has been experimentally observed, for instance as an extremely low minimum in the cross section of electrons scattering on noble gas atoms (Xe, Kr, Ar) at energies $E \simeq 0.7$ eV.

Example 13. Low-energy scattering on a hard-sphere potential. Consider a hard-sphere potential with range R, such that the wavefunction $u_0(r) = 0$ for $r \leq R$ while it is a free particle $u_0(r) \propto \sin(kr - kR)$ for $r > R$. As required by continuity, we see that $u_o(r = R) = 0$. The phase-shift $\delta_0 = -kR$ is thus negative as expected for positive (repulsive) potentials. The total cross section contribution from $l = 0$ valid for *any* energy is then

$$\sigma_0 = \frac{4\pi}{k^2}\sin^2(kR). \tag{6.87}$$

For low energies, the partial wave $l = 0$ gives the dominant contribution. In this case, $k \ll 1/R$ so that $\sin(kR) \simeq kR$, which gives $\sigma \simeq 4\pi R^2$. It is interesting to note that this QM expression is four times as large as the classical limit for this potential, $\sigma_{\text{classical}} = \pi R^2$. What is the physical reason for this? We can understand this result by realizing that in QM, particles have a wave character. Therefore, the particles will probe the entire *surface area* of the hard spheres rather than just their cross-section, similarly to how water waves would interact with an object. For the opposite limit of high energies, one obtains $\sigma = 2\pi R^2$, which still is different from $\sigma_{\text{classical}}$.

Resonant scattering.
To illustrate this phenomenon, consider low-energy scattering on a well-potential:

$$V(r) = \begin{cases} -V_0 \text{ for } r \leq R \\ 0 \text{ for } r > R \end{cases} \tag{6.88}$$

We know by now that at low energies $kR \ll 1$, only the partial wave $l = 0$ contributes significantly to the cross section σ, according to $\sigma = \frac{4\pi}{k^2}\sin^2\delta_0$. The task is to determine δ_0. To do so, we must relate the solution for $r \leq R$, $u_0(r) = A\sin(\kappa r)$ with $\kappa = \sqrt{2m(E + V_0)/\hbar^2}$, with the solution for $r > R$, $u_0(r) = B\sin(kr + \delta_0)$ with $\kappa = \sqrt{2mE/\hbar^2}$. This is accomplished by continuity of u_0 and u_0' (usual boundary conditions) at $r = R$, which yields

$$\tan(\delta_0 + kR) = \frac{k}{\kappa}\tan(\kappa R). \tag{6.89}$$

Now, neglecting the small term kR compared to δ_0, we obtain

$$\sin^2\delta_0 = \frac{\tan^2\delta_0}{1 + \tan^2\delta_0} = \frac{k^2\tan^2(\kappa R)}{\kappa^2 + k^2\tan^2(\kappa R)}. \tag{6.90}$$

The total cross section finally takes the form

$$\sigma = \frac{4\pi}{k^2+\xi^2}, \text{ where } \xi = \frac{\kappa}{\tan(\kappa R)}. \tag{6.91}$$

For a given k, the cross section is maximal when $\xi = 0$, i.e. when $\kappa R = (n+1/2)\pi$ and n is an integer. Inserting the definition of κ, we get:

$$E = -V_0 + \frac{\hbar^2\pi^2}{2mR^2}\Big(n+\frac{1}{2}\Big)^2. \tag{6.92}$$

Physically, this means that when the incident particle has just the right resonant energy satisfying the above equation, it will have a tendency to be bound by the potential and remain at $r \leq R$, thus causing a major disturbance of the wavefunction $\rightarrow$ large σ.

G. The optical theorem

We previously proved the relation $\sigma = \frac{4\pi}{k}\text{Im}\{f(0)\}$ for a spherically symmetric potential $V(r)$. Now, we will demonstrate that this theorem is in fact a direct consequence of particle conservation: for a stationary problem, the net flux of particles into any volume has to equal the net flux out. Choosing the volume as a sphere of radius r centered around the scattering center, this means that $\int j_r r^2 d\Omega = 0$ where j_r is the radial probability current density. We know that this is given by

$$j_r = \text{Re}\{\psi^* \frac{\hbar}{\text{i}m}\partial_r\psi\}. \tag{6.93}$$

To compute this, we choose r to be so large that we can use the asymptotic expression $\psi = \text{e}^{\text{i}\boldsymbol{k}\cdot\boldsymbol{r}} + f(\theta,\phi)\text{e}^{\text{i}kr}/r$.

We begin by computing

$$r^2\psi^*\partial_r\psi = \mathrm{i}kr^2\cos\theta + \mathrm{i}kr\mathrm{e}^{-\mathrm{i}kr(1-\cos\theta)}\cos\theta f^* + (\mathrm{i}kr-1)\mathrm{e}^{\mathrm{i}kr(1-\cos\theta)}f + (\mathrm{i}k-1/r)|f|^2. \tag{6.94}$$

By now multiplying with $\hbar/\mathrm{i}m$, taking the real part of the expression and integrating over all angles, we should obtain zero according to $\int j_r r^2 d\Omega = 0$. Now, the last term ($\frac{\hbar}{\mathrm{i}mr}|f|^2$) in Eq. (6.94) becomes purely imaginary and gives no contribution. The first term $\propto \cos\theta$ gives zero upon integration since $\int_0^\pi \cos\theta\sin\theta = 0$. Finally, we also get rid off the term $\propto k|f|^2$ by using that $\int |f|^2 d\Omega = \sigma$. Dividing the remaining terms in the equation by $\hbar k/m$ gives:

$$\mathrm{Re}\Big\{\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}[r\mathrm{e}^{-\mathrm{i}kr(1-\cos\theta)}\cos\theta f^* + (r+\mathrm{i}/k)\mathrm{e}^{\mathrm{i}kr(1-\cos\theta)}f]d\phi\sin\theta d\theta\Big\} + \sigma = 0. \tag{6.95}$$

Introducing $\cos\theta = x$ and using that $1 + \mathrm{i}/kr \simeq 1$ for large r, this equation becomes

$$\sigma = -r\mathrm{Re}\Big\{\int_0^{2\pi}\int_{-1}^{1}[\mathrm{e}^{-\mathrm{i}kr+\mathrm{i}krx}f^* + \mathrm{e}^{\mathrm{i}kr-\mathrm{i}krx}f]dx\, d\phi\Big\}. \tag{6.96}$$

We now make use of a previously derived result, namely Eq. (6.73). This expansion can be used with $s = kr$ in the first term of the l.h.s. in Eq. (6.96) and $s = -kr$ for the second term. The result is

$$\sigma = \frac{1}{k}\int_0^{2\pi} 2\mathrm{Im}f(0,\phi)d\phi = \frac{4\pi}{k}\mathrm{Im}f(0). \tag{6.97}$$

We here used that in the forward-scattering direction $\theta = 0$, there can be no ϕ-dependence → ϕ-integration merely gives a factor 2π. We have thus proven the optical theorem. We note that:

1. In contrast to our previous derivation using the method of partial waves, we now did not maky any assumption about the potential being spherically symmetry.

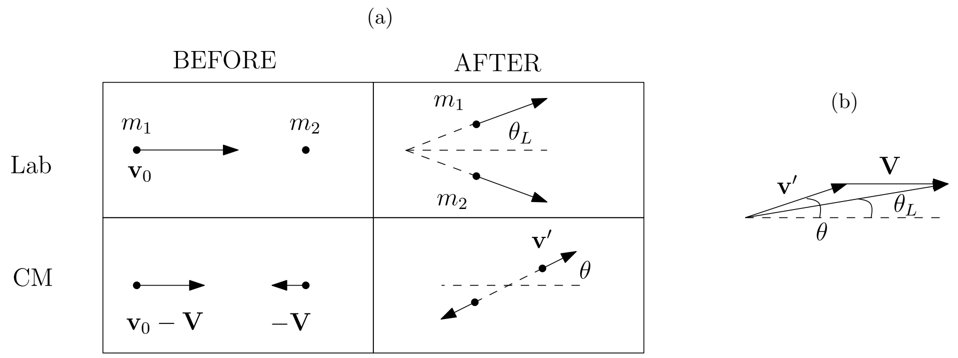
2. The expression we found in the Born-approximation for a spherically symmetric potential, $f^B(\theta) = -\frac{2m}{\hbar q^2}\int_0^\infty V(r)\sin(qr)rdr$ is real. This means that it cannot be used in the optical theorem, since it gives $\sigma = 0$.

3. We have assumed elastic scattering. If inelastic processes occur, *e.g.* exciting internal degrees of freedom in the particle or fragmentation of particles, the net current through a volume is no longer zero. Instead, it must be *negative* since inelastic scattering processes remove particles from their original state. In this more general case, the optical theorem reads

$$\sigma_{\text{el}} + \sigma_{\text{inel}} = \frac{4\pi}{k}\mathrm{Im}\{f_{\text{el}}(0)\}. \tag{6.98}$$

H. Lab- and CM-system

In our scattering theory so far, we have considered a particle scattering on a stationary potential. In fact, this corresponds to the center of mass (CM) frame of a two-particle problem with potential $V(\boldsymbol{r}_1 - \boldsymbol{r}_2)$ since such a scenario can be reduced to an effective one-body problem. We now want to analyze the difference between the lab and CM frames for two particles scattering off each other, defined in the figure below.



Consider the lab-frame where particle 1 with mass m_1 and velocity v_0 scatters on mass m_2 with zero velocity. The CM velocity in the lab-frame is then:

$$V = \frac{m_1 v_0}{m_1 + m_2}. \tag{6.99}$$

We seek the relation between the scattering angles θ and θ_L in the CM and lab frame, respectively. The above figure shows that:

$$\tan\theta_L = \frac{v'\sin\theta}{v'\cos\theta + V}. \tag{6.100}$$

Note that $v' = |\boldsymbol{v}'| = v_0 - V$ (magnitude of velocity of the scattered and incident particle in the CM frame is the same). Inserting v' and V into the above expression for $\tan\theta_L$ yields:

$$\tan\theta_L = \frac{\sin\theta}{\cos\theta + m_1/m_2}. \tag{6.101}$$

Thus, when the target mass $m_2 \to \infty$, we obtain $\theta_L = \theta$, as expected. Moreover, for equal masses $m_1 = m_2$, one obtains $\tan\theta_L = \tan\theta/2$, so that $\theta_L = \theta/2$ is maximally $\pi/2$.

To identify a relation between the scattering cross sections in the two frames, we will need:

$$\cos\theta_L = \frac{1}{\sqrt{1+\tan^2\theta_L}} = \frac{\cos\theta + \gamma}{\sqrt{1 + 2\gamma\cos\theta + \gamma^2}} \tag{6.102}$$

where we defined $\gamma \equiv m_1/m_2$. Moreover, the relation between the solid angles is:

$$\frac{d\Omega_L}{d\Omega} = \frac{\sin\theta_L d\theta_L}{\sin\theta d\theta} = \frac{d\cos\theta_L}{d\cos\theta} = \frac{1 + \gamma\cos\theta}{(1 + 2\gamma\cos\theta + \gamma^2)^{3/2}}. \tag{6.103}$$

Now, the particle flux incident toward the target should only depend on the relative velocity between 1 and 2 in both systems and is thus the same. Also, the same number of particles have to be scattered into $d\Omega$ and $d\Omega_L$: the physics cannot be different by changing referenc frame. Because of the above two facts, it follows from the definition of the differential scattering cross section that

$$d\sigma_L(\theta_L, \phi) = d\sigma(\theta, \phi). \tag{6.104}$$

Hence, we find that

$$\frac{d\sigma_L}{d\Omega_L} = \frac{d\sigma}{d\Omega}\frac{(1 + 2\gamma\cos\theta + \gamma^2)^{5/2}}{1 + \gamma\cos\theta}. \tag{6.105}$$

I. Scattering of identical particles

We now consider what happens when two identical particles are scattered on each other. It is known that the two-particle state satisfies $\psi(1,2) = \psi(2,1)$ for bosons and $\psi(1,2) = -\psi(2,1)$ for fermions where $1 = (\boldsymbol{r}_1, s_1)$ and $2 = (\boldsymbol{r}_2, s_2)$. Consider only the spatial part of the wavefunction to begin with and focus on the CM frame. In this case, a two-particle state ψ that is symmetric (antisymmetric) in $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ must be an even (odd) function of the relative-coordinate $\boldsymbol{r} \equiv \boldsymbol{r}_1 - \boldsymbol{r}_2$. In spherical coordinates, $\boldsymbol{r} \to -\boldsymbol{r}$ means that $(r, \theta, \phi) \to (r, \pi - \theta, \phi + \pi)$. Now, our asymptotic wavefunction is neither symmetric nor antisymmetric in the form that we have used it. Therefore, for identical particles it must be replaced with

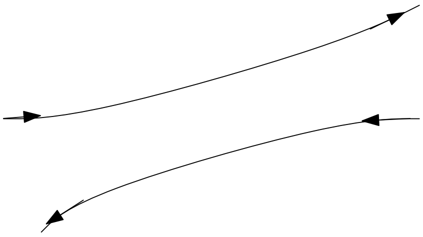
$$\psi(1,2) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \pm \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} + [f(\theta) \pm f(\pi - \theta)]\mathrm{e}^{\mathrm{i}kr}/r. \tag{6.106}$$

The upper sign is used for a symmetric wavefunction, and the lower for an antisymmetric wavefunction. Note how the spherical part $\mathrm{e}^{\mathrm{i}kr}/r$ accounts for scattering of particlecs in diamatrically opposite directions. As before, $d\sigma$ is defined by the ratio of the particle flux into $\mathrm{d}\Omega$ and the incident particle stream for one of two plane waves:

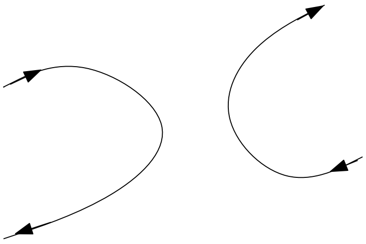
$$\frac{d\sigma}{d\Omega} = |f(\theta) \pm f(\pi - \theta)|^2. \tag{6.107}$$

It makes sense physically that the scattering of *both* particles must be taken into account when they are identical, because we cannot distinguish between the following scenarios shown in the figure.

(a)



(b)



Eq. (6.107) is consistent with the standard QM treatment: add "wavefunctions", then take absolute value squared to compute probabilities. Note that this is different from how we would classically allow for the two possibilities: $d_\Omega\sigma = |f(\theta)|^2 + |f(\pi-\theta)|^2$, which has no interference term between $f(\theta)$ and $f(\pi-\theta)$. Whether or not to use symmetric or antisymmetric states ψ depends on the spin configuration. We now proceed to illustrate this.

Scattering of spin-0 particles.

Spin-0 particles: bosons → spatially symmetric wavefunction. We thus use the upper sign in $d_\Omega\sigma$. Assume for concreteness that the bosons interact via the Coulomb-potential, for which case we have

$$f_C(\theta) = \frac{n}{2k\sin^2(\theta/2)}\mathrm{e}^{-2\mathrm{i}\ln\sin(\theta/2)+\mathrm{i}\delta}. \tag{6.108}$$

Here, $n = Z^2m/(ka_0m_e)$. Inserting this into $d_\Omega\sigma$:

$$\frac{d\sigma}{d\Omega} = \Big(\frac{Z^2e^2}{4\pi\epsilon_0E_T}\Big)^2\Big[\sin^{-4}(\theta/2) + \cos^{-4}(\theta/2) + \frac{2\cos[n\,\ln\tan^2(\theta/2)]}{\sin^2(\theta/2)\cos^2(\theta/2)}\Big] \tag{6.109}$$

where $E_T = \hbar^2k^2/2m$. The last term is a purely QM effect stemming from the interference between $f(\theta)$ and $f(\pi-\theta)$. This effect due to identical particles in QM has been verified experimentally for C^{12} scattering on carbon [see Phys. Rev. Lett. **4**, 365 (1960)].

Scattering of particles with spin.

Even if the interaction between two spinful particles does not depend on the spin itself, we must consider the fact that the particles have spin to obtain the correct $d_\Omega\sigma$. To see this, consider $e-e$ scattering (spin 1/2). Now, two spin 1/2 states may be combined into one singlet ($\uparrow\downarrow - \downarrow\uparrow$) or three triplet ($\uparrow\uparrow, \downarrow\downarrow, \uparrow\downarrow + \downarrow\uparrow$) states. If the particles are randomly polarized: probability 1/4 for singlet and probability 3/4 for triplet state. This yields:

$$d_\Omega\sigma = \frac{3}{4}|f(\theta) - f(\pi-\theta)|^2 + \frac{1}{4}|f(\theta) + f(\pi-\theta)|^2. \tag{6.110}$$

From this, we can infer that particles scattered into $\theta = \pi/2$ must be singlets, since the triplet contribution is zero for this angle. Moreover, if the spins are not initially random, but fully polarized in the same direction (i.e. triplets), there can be no scattering into $\theta = \pi/2$.

VII. MAGNETIC FIELDS IN QUANTUM MECHANICS

Learning goals. After reading this chapter, the student should:

- Be able to write down how the presence of a magnetic field, and thus vector potential, can modify the Hamiltonian of a system, and what the physical meaning of each corresponding term is.
- Be able to explain what Landau levels are, how they behave physically, and what type of physical consequences they lead to.
- Be able to explain what the Aharanov-Bohm effect is and to mathematically outline the basic equations describing it.

We shall consider here the influence of magnetic fields on QM systems. This is extremely important because it is one of the simplest and most common experimental ways to manipulate eigenfunctions and energy levels.

A. Zeeman effect

Normal Zeeman effect.
To incorporate a $\boldsymbol{B}$-field into the SE, we know from earlier treatment (*e.g.* classical mechanics) that we should include a gauge field $\boldsymbol{A}$. For $\boldsymbol{B} = B\hat{\boldsymbol{z}}$, we may use *e.g.* $\boldsymbol{A} = \frac{B}{2}(-y, x, 0)$. The difference between Hamiltonians with and without a magnetic field becomes

$$\begin{aligned}\hat{H} - \hat{H}_0 &= \frac{(\hat{\boldsymbol{p}} - q\boldsymbol{A})^2}{2m} - \frac{\hat{\boldsymbol{p}}^2}{2m} = -\frac{q}{m}\boldsymbol{A}\cdot\hat{\boldsymbol{p}} + \frac{q^2}{2m}\boldsymbol{A}^2 \\ &= -\frac{qB}{2m}(x\hat{p}_y - y\hat{p}_x) + \frac{q^2B^2}{8m}(x^2+y^2). \end{aligned} \tag{7.1}$$

Since $xp_y - yp_x = L_z$ is the angular momentum component in the z-direction, we may write

$$\hat{H}' \equiv \hat{H} - \hat{H}_0 = -\frac{q}{2m}B\hat{L}_z + \frac{q^2B^2}{8m}(x^2+y^2) = \hat{H}_1' + \hat{H}_2'. \tag{7.2}$$

Focus now on the term $\hat{H}_1' \equiv -\boldsymbol{\mu}_L \cdot \boldsymbol{B}$ since $\hat{H}_2'$ is quadratic and negligible for small B. We treat $\hat{H}_1'$ as a perturbation and we defined $\boldsymbol{\mu}_L = q\boldsymbol{L}/2m$. $\hat{H}_1'$ describes the coupling between external field and the induced field of a charged particle with orbital angular momentum.

For a spherically symmetric potential, $R_{nl}(r)Y_{lm}(\theta,\phi)$ are eigenfunctions for $\hat{L}_z$ with eigenvalue $\hbar m$. Considering an electron ($q = -e$ and $m = m_e$), the added energy due to $\hat{H}_1'$ becomes

$$\Delta E = \frac{e\hbar B}{2m_e}m = \mu_B B m \;(m = -l, \ldots, l). \tag{7.3}$$

We defined the Bohr-magneton $\mu_B = e\hbar/2m_e$. Every energy level is thus split into $2l+1$ levels with a spacing depending on B and not on the quantum numbers n or l. This is the normal Zeeman effect, but when we take into account spin we obtain the experimentally observed anomalous Zeeman effect.

Anomalous Zeeman effect.
A particle with spin has an additional *internal* angular momentum $\boldsymbol{\mu}_s$ which also couples to the magnetic field. The total perturbation then becomes $\hat{H}_1' = \frac{eB}{2m_e}(\hat{L}_z + 2\hat{S}_z)$ where we used $g_s = 2$ as the Lande g-factor. However, we must also consider how spin influences $\hat{H}_0$, i.e. the B-independent part. This part gains a spin-orbit interaction

$$\hat{H}_{\text{so}} = f(r)\boldsymbol{L}\cdot\boldsymbol{S} \tag{7.4}$$

so that the eigenfunctions now depend on $\boldsymbol{J}^2$ and J_z where $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, as $\boldsymbol{L}$ and $\boldsymbol{S}$ are no longer conserved separately (the Hamiltonian does not commute with either in the presence of $\hat{H}_{\text{so}}$). We proceed to distinguish beween weak and strong magnetic fields.

Strong fields: In this case, we may disregard $\hat{H}_{\text{so}}$ relative the magnetic term $\hat{H}_1'$. The resulting splitting is then simply

$$\Delta E = \frac{e\hbar}{2m_e}(m + 2m_s)B = \mu_B B(m + 2m_s). \tag{7.5}$$

A given energy level is then split into $2l + 3$ levels for $l > 0$ since $m + 2m_s$ takes values between $-l - 1$ and $l + 1$. For $l = 0$, the splitting is into two levels.

Weak fields: This is a more complicated situation since we cannot disregard $\hat{H}_{\text{so}}$ anymore. Thus, the eigenstates of the unperturbed Hamiltonian are $|j, m_j, l\rangle$. This is because $\hat{H}_0 + \hat{H}_{\text{so}}$ commutes with both $\boldsymbol{J}^2, J_z$, and $\boldsymbol{L}^2$. Thus, the perturbation energy becomes:

$$\Delta E = \mu_B B\langle j, m_j, l|L_z + 2S_z|j, m_j, |l\rangle = mu_B B(\hbar m_j + \langle j, m_j, l|S_z|j, m_j, l\rangle). \tag{7.6}$$

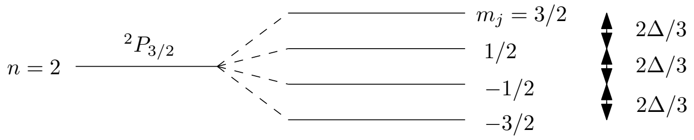
Here, j and m_j are the quantum numbers determining the eigenvalues for the operators $\boldsymbol{J}^2$ and J_z. For $s = 1/2$, there are two possible values for j: $j = l \pm 1/2$. To compute the expectation value of S_z, we want to express $|j.m_j, l\rangle$ in terms of eigenspinors for S_z. This is a length but straightforward calculation which we do not show here (see introductory QM course and angular momentum operator algebra), but simply state the final result:

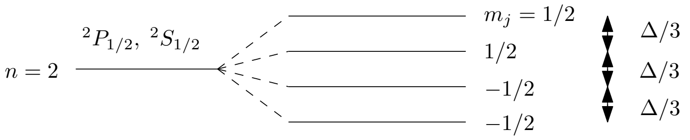
$$\langle j, m_j, l|S_z|j, m_j, l\rangle = \pm\frac{\hbar m_j}{2l + 1}. \tag{7.7}$$

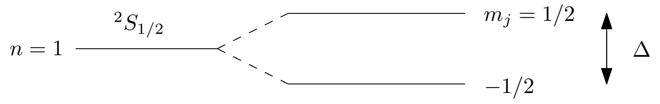
Inserted into ΔE, we obtain:

$$\Delta E = \mu_B B\frac{2j + 1}{2l + 1}m_j,\ (j = l \pm 1/2, m_j = -j, \ldots, j). \tag{7.8}$$

This gives rise to a different energy splitting with a spacing that is no longer independent on the quantum numbers. We show the magnetic field splitting for the hydrogen $n = 1$ and $n = 2$ levels in the figure below.







We have introduced the notation $^{2S+1}L_J$ to characterize the levels, where S is the quantum number for total spin (1/2 in our case), L is the quantum number for total orbital angular momentum $(S : l = 0, P : l = 1, D : l = 2, \ldots)$, while J is the quantum number for total angular momentum $\boldsymbol{J}$. Moreover, $\Delta \equiv 2\mu_B B$. For instance,$^2P_{1/2}$ then means $s = 1/2, l = 1, j = 1/2$.

B. Landau levels

The Zeeman effect is concerned with the effect of $\boldsymbol{B}$ on bound electrons, such as the coupling between spin $\boldsymbol{S}$ and field $\boldsymbol{B}$. We consider free electrons, neglecting spin for now, and show that for a constant $\boldsymbol{B}$, the SE can be solved exactly. We use a Landau-gauge $\boldsymbol{A} = (-By, 0, 0)$ so that the Hamiltonian for $q = -e$ becomes:

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \frac{\mathrm{i}e\hbar B}{m}y\partial_x + \frac{e^2B^2}{2m}y^2. \tag{7.9}$$

This $\hat{H}$ commutes with $\hat{p}_x$ and $\hat{p}_z$ and thus admits common eigenstates with these operators. The general solution should then have the form $\psi(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}k_x x + \mathrm{i}k_z z}\phi(y)$. Inserted into $\hat{H}\psi = E\psi$, we obtain the following equation for ϕ:

$$-\frac{\hbar^2}{2m}\phi'' + \left[\frac{\hbar^2 k_x^2}{2m} - \frac{\hbar e B k_x}{m}y + \frac{e^2B^2}{2m}y^2\right]\phi + \frac{\hbar^2 k_z^2}{2m}\phi = E\phi. \tag{7.10}$$

This can be written in a more compact manner:

$$-\frac{\hbar^2}{2m}\phi'' + \frac{1}{2}m\omega_c^2(y-y_0)^2\phi = \tilde{E}\phi. \tag{7.11}$$

where $E = \tilde{E} + \hbar^2k_z^2/2m$, $y_0 = \hbar k_x/eB$, and $\omega_c = eB/m$. We see that ω_c is the cyclotron frequencey: classical angular frequency for the circular motion of an electron in a $\boldsymbol{B}$ field. Now, Eq. (7.11) has a familiar form: a harmonic oscillator centered around y_0. We immediately know what the eigenvalues are according to our detailed previous treatment of such a system:

$$\tilde{E} = (n+1/2)\hbar\omega_c \to E = (n+1/2)\hbar\omega_c + \hbar^2k_z^2/2m \; (n = 0, 1, 2, \ldots). \tag{7.12}$$

The belonging eigenfunctions are:

$$\psi(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}k_x x + \mathrm{i}k_z z}\phi_n(y - y_0) \tag{7.13}$$

where ϕ_n is the n-th harmonic oscillator function. The energy E for our particle thus has two parts: free particle motion along the $\boldsymbol{B}$ field (z-axis) and quantized motion perpendicularly to $\boldsymbol{B}$ (xy-plane).

Landau-levels: quantized levels for fixed k_z (varying n).

Landau-bands: continuous energy bands for fixed n (varying k_z).

An important aspect is that since the energy does not depend on the quantity $y_0 \propto k_x$, the Landau levels are *massively degenerate*. To see this, consider a large but finite volume $V = L_x L_y L_z$. Using periodic boundary conditions, *e.g.* $\psi(x + L_x) = \psi(x)$, the allowed values of the momenta are $k_x = 2\pi n_x/L_x$ and $k_z = 2\pi n_z/L_z$ where n_i are integers. Note that using periodic boundary conditions allows us to use free-particle wavefunctions to count the number of states in contrast to hard-wall boundary conditions where $\psi = 0$ at the edges, while still obtaining the same density of states. In turn, this means that the allowed values for $y_0 = \hbar k_x/eB$ are separated by $\Delta y_0 = h/eBL_x$. The number of available positions for y_0 then becomes:

$$\frac{L_y}{\Delta y_0} = L_x L_y \frac{eB}{h} = \frac{\Phi_{\text{tot}}}{h/e}. \tag{7.14}$$

Here, $\Phi_{\text{tot}} = L_x L_y B$ is the total magnetic flux through the area $L_x \times L_y$. We may conclude that each Landau level contains the same number of states: $\Phi_{\text{tot}}/(h/e)$. Each state then carries a flux quantum $\Phi_0 = h/e$.

Oscillation of the Fermi level.
We saw above that the degree of degeneracy of Landau levels was $L_x L_y Be/h$ per level. Taking spin into account, the degeneracy is doubled.

Thus, if the 2D electron density of the system is n_2, meaning there are in total $n_2 L_x L_y$ electrons, they can all reside in the same Landau level if the field is so strong that

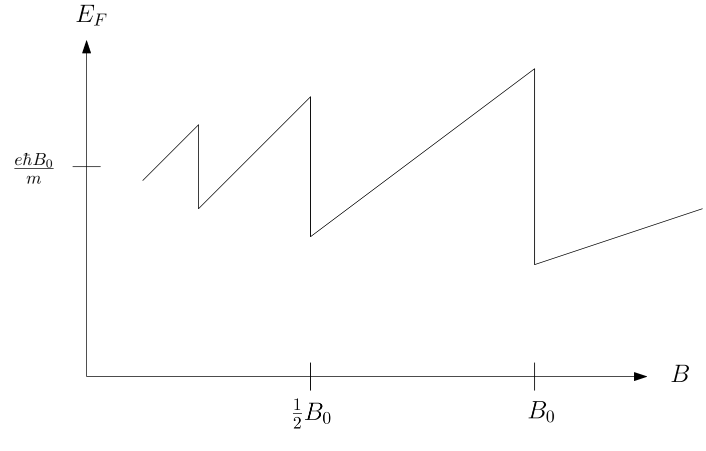
$$B > B_0 = \frac{1}{2} n_2 \frac{h}{e}. \tag{7.15}$$

Consider in fact a 2D electron gas, which is typically studied experimentally in the context of Landau levels. We thus disregard excitations in the z-direction. Let us compute the Fermi energy E_F as a function of B. Note that: both the degeneracy of Landau levels (LL) and the Landau level energy itself $[E_n = (n + 1/2)\hbar eB/m]$ are proportional to B.

- For $B > B_0$, all electrons are in the lowest LL so that $E_F = \frac{1}{2} e\hbar B/m$.
- If $B_0/2 < B < B_0$, the electrons that can't fit into the lowest LL have room to be in the second lowest LL: $E_F = 3e\hbar B/2m$.

This argument is repeated as B decreases. At $B = B_0$, E_F jumps from $\frac{1}{2} e\hbar B_0/m$ to $\frac{3}{2} e\hbar B_0/m$. At $B = \frac{1}{2} B_0$, E_F jumps from $\frac{3}{4} e\hbar B_0/m$ to $\frac{5}{4} e\hbar B_0/m$, et.c.

In general: discontinuities at $B = B_0/k$ where E_F jumps between $(1 - \frac{1}{2k})e\hbar B_0/m$ and $(1 + \frac{1}{2k})e\hbar B_0/m$.



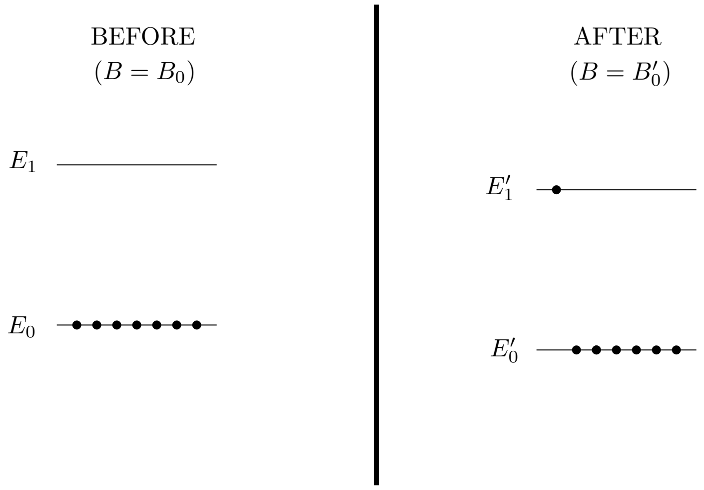
For a 3D system, this picture is slightly modified. When $B < B_0$, the electrons that don't fit into the 0th LL will not directly go into the 1st LL, but instead populate states $k_z \neq 0$ with energy

$$E = \frac{1}{2}\hbar\omega_c + \frac{\hbar^2 k_z^2}{2m} \tag{7.16}$$

This energy will increase until it equals the energy of the 1st LL, and then this level is starting to fill up. We still have a sharp peak in E_F vs. B every time a new LL is activated. Since the transport properties of metals are determined by the electrons at the Fermi level, the strong variation of E_F vs. B is manifested *e.g.* in the conductivity ρ which oscillates with B. This is the Shubnikov-de Haas effect.

Total energy as a function of magnetic field.
When B is lowered just below B_0, one electron jumps up to E_1 and we now know that E_F should make a jump. But what happens to the total system energy? Naively, one might first think that since energy is supplied by the external field, we should simply see a decrease in E_{tot} when B is reduced. However, it turns out that the physics is a bit more interesting than that.

Consider a system of size $L^2 = L_xL_y$ and we have N electrons in total. We have seen that the degeneracy of a LL is $2L^2B_0e/h$ when spin is taken into account. Thus, all e^- in our system fit into E_0 when $2L^2B_0e/h = N$ meaning that $B_0 = Nh/2L^2e$. Thus, we have the situation shown in the figure when we start out with B_0 and then lower the field so that one e^- jumps up to E_1.



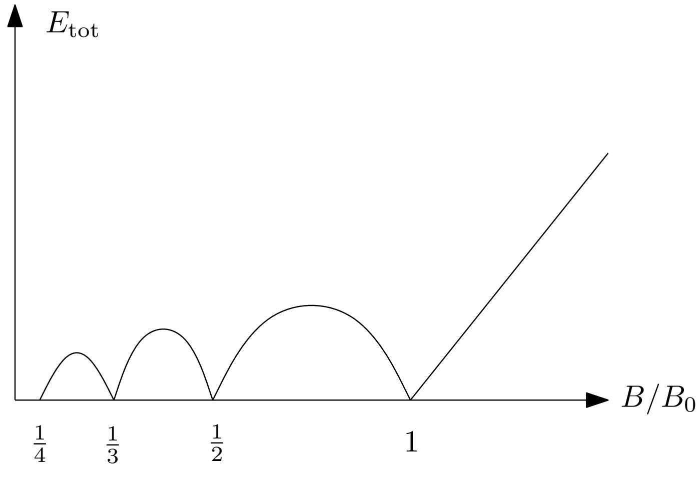
The question is now: what is the change in the total energy of the system, in effect $\Delta E = E_{\text{tot}} - E'_{\text{tot}}$? We know that $E_0(B) = \hbar eB/2m$ and $E_1(B) = 3\hbar eB/2m$. Moreover, the field B_0 that forces one electron to leave the lowest LL must by definition satisfy

$$\frac{2L^2B_0'e}{h} = N - 1 \to B_0' = B_0 - \frac{h}{2L^2e} < B_0. \tag{7.17}$$

Now, we may evaluate ΔE:

$$\Delta E = NE_0 - [(N-1)E_0' + E_1'] = \frac{\pi\hbar^2}{2L^2m}(2 - N). \tag{7.18}$$

The energy thus increases if $N > 2$. If $N = 2$, there is no change since $B_0' = B_0/2$. For large N, the total energy will in general oscillate as shown in the figure.



So energy decreases, but non-monotonically. Note that this picture changes if we account for the Zeeman-splitting of the electrons, since it removes the factor 2 in the spin degeneracy of the states.

C. Aharonov-Bohm effect

Wavefunction in space with $\boldsymbol{B} = 0$.
Assume that $\boldsymbol{B}(\boldsymbol{r}) \neq 0$ is present in some region of space, whereas other regions have $\boldsymbol{B} = 0$. An example is a very long coil with a current running through it, which to a very good approximation only has $\boldsymbol{B} \neq 0$ inside it. In the regions where $\boldsymbol{B} = 0$, we have $\boldsymbol{A} = \nabla\lambda$ since $\nabla \times \boldsymbol{A} = 0$ there. This means that (up to a constant):

$$\lambda(\boldsymbol{r}) = \int_{\boldsymbol{r}_0}^{\boldsymbol{r}} \boldsymbol{A}(\boldsymbol{s}) \cdot d\boldsymbol{s} \tag{7.19}$$

where $\boldsymbol{r}_0$ is an arbitrary point in the region where $\boldsymbol{B} = 0$. The integration path is arbitrary, as long as we stay inside the $\boldsymbol{B} = 0$ region. The wavefunction is obtained from the SE:

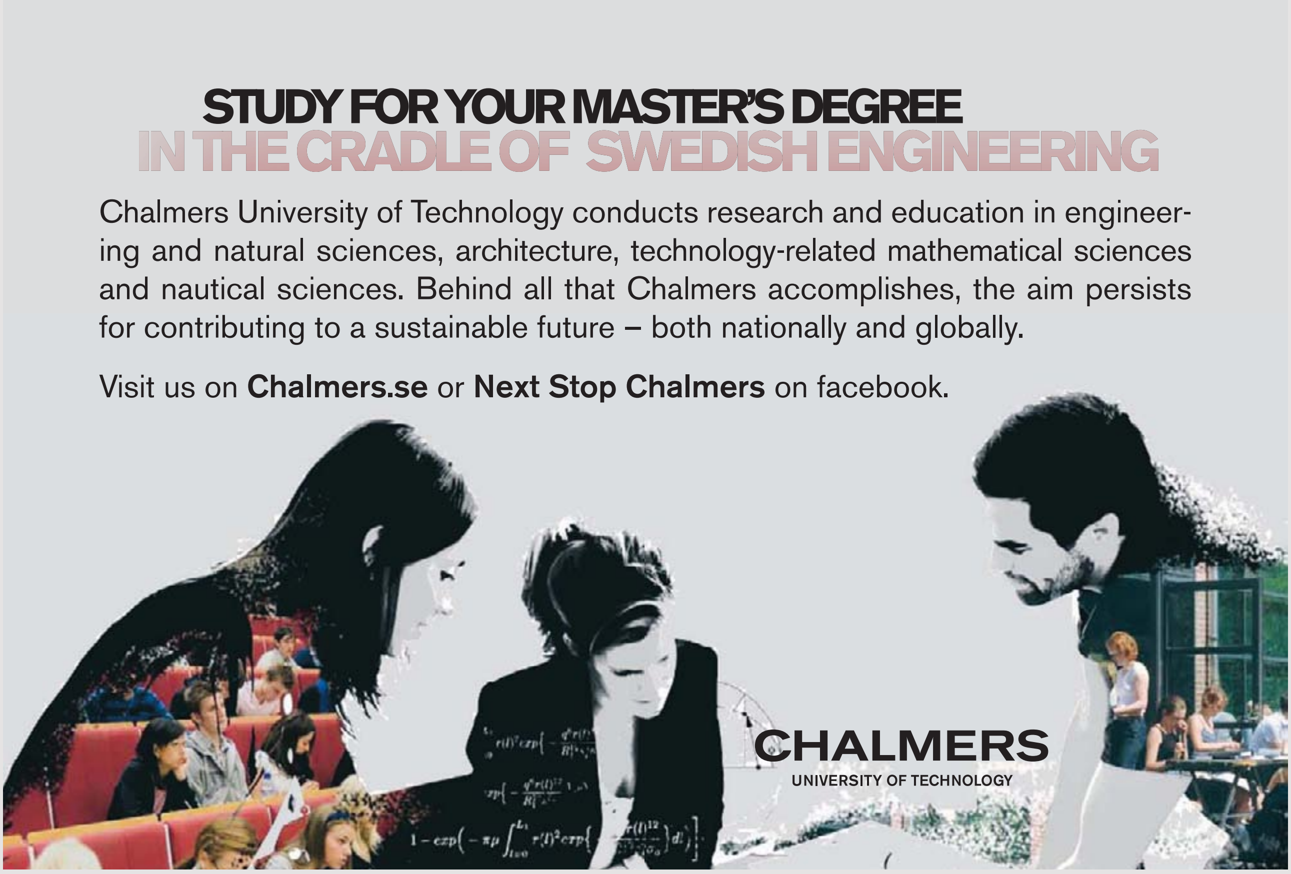
$$\mathrm{i}\hbar\partial_t\psi = \frac{1}{2m}\Big(\frac{\hbar}{\mathrm{i}}\nabla - q\boldsymbol{A}\Big)^2\psi + V(\boldsymbol{r})\psi \tag{7.20}$$

where $V(\boldsymbol{r})$ is potential energy stemming from other effects than the field. The physics must be gauge-invariant. Thus, let us perform a gauge-transformation:

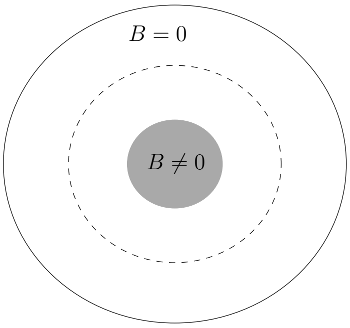
$$\boldsymbol{A}' = \boldsymbol{A} + \nabla\chi,\ \phi' = \phi - \partial_t\chi,\ \psi' = \psi\mathrm{e}^{\mathrm{i}q\chi/\hbar}. \tag{7.21}$$

Now choose $\chi = -\lambda$ so that:

$$\mathrm{i}\hbar\partial_t\psi' = \frac{1}{2m}\Big(\frac{\hbar}{\mathrm{i}}\nabla\Big)^2\psi' + V(\boldsymbol{r})\psi'. \tag{7.22}$$



This is the same equation as if $\boldsymbol{A}=0$ in the first place. *However, we cannot in general just set $\boldsymbol{A}=0$ in a region where $\boldsymbol{B}=0$ if the field-free area encloses a region where $\boldsymbol{B}\neq 0$.* To see this, consider the geometry shown in the figure.

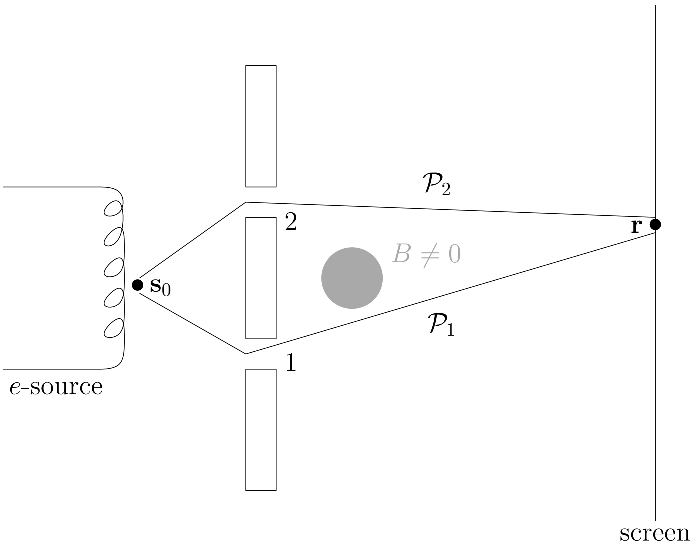


If we integrate along the dashed line, we obtain

$$\oint \boldsymbol{A}(\boldsymbol{s})\cdot d\boldsymbol{s} = \int\int_S \nabla\times\boldsymbol{A}\, dS = \Phi_S. \tag{7.23}$$

where S is the shaded region and Φ_S is the flux through it. This equation shows that $\boldsymbol{A}$ cannot be zero everywhere in the white region where $\boldsymbol{B}=0$.

Interference experiment.
Does this mean that an electron moving only in an area with $\boldsymbol{B}=0$ can still be affected by the presence of $\boldsymbol{B}\neq 0$ in the inaccessible region? Aharanov and Bohm suggested the following interference experiment to clarify this. The figure shows an electron source that emits e^- from the point $\boldsymbol{s}_0$, and the electrons consequently pass through the slits 1 and 2 and hit the screen at $\boldsymbol{r}_1$ having taken the paths $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively.



Assume that the shaded region is completely inaccessible to the electrons. The total wavefunction is a superposition of the contribution from paths $\mathcal{P}_1$ and $\mathcal{P}_2$:

$$\psi_{\text{tot}} = \psi_{\mathcal{P}_1}(\boldsymbol{r},t) + \psi_{\mathcal{P}_2}(\boldsymbol{r},t). \tag{7.24}$$

According to our previous treatment, we have:

$$\begin{aligned}\psi_{\mathcal{P}_1} &= \psi_0(\boldsymbol{r},t)\mathrm{e}^{\mathrm{i}\frac{e}{\hbar}\int_{\mathcal{P}_1}\boldsymbol{A}(\boldsymbol{s})\cdot d\boldsymbol{s}},\\ \psi_{\mathcal{P}_2} &= \psi_0(\boldsymbol{r},t)\mathrm{e}^{\mathrm{i}\frac{e}{\hbar}\int_{\mathcal{P}_2}\boldsymbol{A}(\boldsymbol{s})\cdot d\boldsymbol{s}}.\end{aligned} \tag{7.25}$$

$\psi_0(\boldsymbol{r},t)$ is as before the wavefunction for $\Phi=0$. Note that the relative phase between $\psi_{\mathcal{P}_1}$ and $\psi_{\mathcal{P}_2}$ is:

$$\int_{\mathcal{P}_1}\boldsymbol{A}(\boldsymbol{s})\cdot d\boldsymbol{s} - \int_{\mathcal{P}_2}\boldsymbol{A}(\boldsymbol{s})\cdot d\boldsymbol{s} = \oint \boldsymbol{A}(\boldsymbol{s})\cdot d\boldsymbol{s} = \Phi. \tag{7.26}$$

Here, Φ is the flux through the shaded region. We may then rewrite

$$\psi_{\text{tot}}(\boldsymbol{r},t) = (\psi_0 \mathrm{e}^{\mathrm{i}e\Phi/\hbar} + \psi_0)\mathrm{e}^{\mathrm{i}\frac{e}{\hbar}\int_{\mathcal{P}_2}\boldsymbol{A}(\boldsymbol{s})\cdot d\boldsymbol{s}}, \tag{7.27}$$

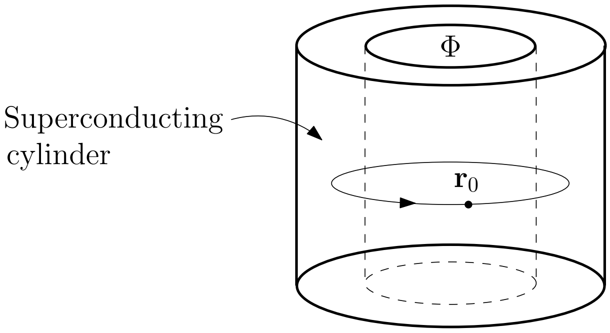
so that the probability density of electrons hitting the screen becomes

$$|\psi_{\text{tot}}|^2 = |\psi_0 \mathrm{e}^{\mathrm{i}e\Phi/\hbar} + \psi_0|^2. \tag{7.28}$$

This means that the interference pattern *changes with* Φ even if the electrons never move in the shaded region! This is the Aharonov-Bohm effect, measured in 1960 by Chambers. We learn therefore that $\boldsymbol{A}$ plays a fundamental role. However, note that the physically measurable quantity $|\psi_{\text{tot}}|^2$ is gauge-independent (only depends on Φ and not $\boldsymbol{A}$).

D. Flux quantization in superconductors

An interesting case where electrons move in field-free space is in a superconductor, which besides having zero electrical resistance also expels $\boldsymbol{B}$ from its interior. For a superconducting cylinder, a flux Φ can pass through the hollow middle, meaning again that $\boldsymbol{A}$ cannot be zero in the superconductor despite $\boldsymbol{B} = 0$ there.



The wavefunction ψ inside the superconductor can again be expressed as the $\Phi = 0$ wavefunction times a phase factor:

$$\psi(\boldsymbol{r}) = \psi_0(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\frac{q}{\hbar}\int_{\boldsymbol{r}_0}^{\boldsymbol{r}}\boldsymbol{A}(\boldsymbol{s})\cdot d\boldsymbol{s}}. \tag{7.29}$$

If the integral path now is taken to form a closed loop inside the superconductor (as shown in the figure), the wavefunction is multiplied with $\mathrm{e}^{-\mathrm{i}q\Phi/\hbar}$. Since the wavefunction has to be single-valued, it follows that

$$\mathrm{e}^{-\mathrm{i}q\Phi/\hbar} = \mathrm{e}^{2\mathrm{i}n\pi},\ n = 0, \pm 1, \pm 2, \dots \tag{7.30}$$

so that

$$\boxed{\Phi = \frac{2\pi\hbar}{q}n.}$$

The flux has to be quantized. This has been experimentally observed and one found the flux to be quantized in units of $\pi\hbar/e$. This corresponds to $q = -2e$, suggesting that in superconductors the fundamental entity is an electron *pair* (in accordance with so-called BCS theory).

VIII. QUANTIZED RADIATION THEORY

Learning goals. After reading this chapter, the student should:

- Be able to qualitatively describe how to quantize the electromagnetic field quantum mechanically.
- Be able to explain the physical significance of coherent states for the radiation field modes.
- Be able to schematically write down the state vector for a fully quantized radiation theory and explain what spontaneous and stimulated emission means.

We have treated the EM field classically so far. However, the EM field is also governed by QM and we now want to treat both the atomic system *and* the field quantum mechanically.

A. Quantization of the radiation field

The starting point for determining the QM Hamilton-operator of a system is to know the classical Hamiltonian. Hence, that is where our investigation begins.

Classical Hamiltonian for the field.
Consider the existence of an EM field without any source terms (charges and currents) in a cubic volume $V = L^3$ with periodic boundary conditions. V is introduced for convenience so that we can quantify the number of modes in the system.

It is $\boldsymbol{A}$ that couples the atomic system with the EM field. In a Coulomb gauge with $\nabla\cdot\boldsymbol{A}=0$, we may write $\boldsymbol{A}$ as a superposition of plane-waves:

$$\boldsymbol{A}(\boldsymbol{r},t)=\sum_{\boldsymbol{k},\lambda}\boldsymbol{e}_{\boldsymbol{k}\lambda}\sqrt{\frac{\hbar}{2\epsilon_0 Vck}}[\tilde{a}_{\boldsymbol{k}\lambda}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}-\mathrm{i}\omega_k t}+\tilde{a}^*_{\boldsymbol{k}\lambda}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}+\mathrm{i}\omega_k t}], \tag{8.1}$$

where $\omega_k=kc$ and $\tilde{a}_{\boldsymbol{k}\lambda}$ are amplitude factors. Note that the photon $\boldsymbol{A}$ is transversely polarized since $\boldsymbol{e}_{\boldsymbol{k}\lambda}\perp\boldsymbol{k}$, guaranteeing that $\nabla\cdot\boldsymbol{A}=0$. The sum $\sum_{\boldsymbol{k},\lambda}$ goes over discrete wavevectors $\boldsymbol{k}=\frac{2\pi}{L}(n_x,n_y,n_z)$ and over the two polarization vectors $\boldsymbol{e}_{\boldsymbol{k},1}$ and $\boldsymbol{e}_{\boldsymbol{k},2}$ (both $\perp\boldsymbol{k}$). Let $a_{\boldsymbol{k}\lambda}\equiv\tilde{a}_{\boldsymbol{k}\lambda}\mathrm{e}^{-\mathrm{i}\omega_k t}$ for brevity of notation. Now, the energy of the field itself is

$$\mathcal{H}=\frac{1}{2}\int_V[\epsilon_0\boldsymbol{E}^2+\frac{1}{\mu_0}\boldsymbol{B}^2]d\boldsymbol{r}=\frac{\epsilon_0}{2}\int_V[(\partial_t\boldsymbol{A})^2+c^2(\nabla\times\boldsymbol{A})^2]d\boldsymbol{r}. \tag{8.2}$$

Inserting our expression for $\boldsymbol{A}$ gives:

$$\begin{aligned}\mathcal{H}=&-\frac{\hbar}{4V}\sum_{\boldsymbol{k}\lambda}\sum_{\boldsymbol{k}'\lambda'}\sqrt{\omega_k\omega_{k'}}[\boldsymbol{e}_{\boldsymbol{k}\lambda}\cdot\boldsymbol{e}_{\boldsymbol{k}'\lambda'}+(\hat{k}\times\boldsymbol{e}_{\boldsymbol{k}\lambda})\cdot(\hat{\boldsymbol{k}}'\times\boldsymbol{e}_{\boldsymbol{k}'\lambda'})]\\&\times\int_V(a_{\boldsymbol{k}\lambda}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}-a^*_{\boldsymbol{k}\lambda}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}})(a_{\boldsymbol{k}'\lambda'}\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}}-a^*_{\boldsymbol{k}'\lambda'}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}})d\boldsymbol{r}.\end{aligned} \tag{8.3}$$

We introduced $\hat{\boldsymbol{k}}=\boldsymbol{k}/|\boldsymbol{k}|$. Using the vector identity

$$(\boldsymbol{A}\times\boldsymbol{B})(\boldsymbol{C}\times\boldsymbol{D})=(\boldsymbol{A}\cdot\boldsymbol{C})(\boldsymbol{B}\cdot\boldsymbol{D})-(\boldsymbol{A}\cdot\boldsymbol{D})(\boldsymbol{B}\cdot\boldsymbol{C}) \tag{8.4}$$

and

$$\frac{1}{V}\int\mathrm{e}^{\mathrm{i}(\boldsymbol{k}_1-\boldsymbol{k}_2)\cdot\boldsymbol{r}}d\boldsymbol{r}=\begin{cases}1\text{ if }\boldsymbol{k}_1=\boldsymbol{k}_2\\0\text{ if }\boldsymbol{k}_1\neq\boldsymbol{k}_2\end{cases} \tag{8.5}$$

since the allowed $\boldsymbol{k}$-values are such that an integer number of wavelengths fit into each side L of the volume, we see that there is only a contribution to the sum $\sum_{\boldsymbol{k}\boldsymbol{k}'}$ from $\boldsymbol{k}'=\pm\boldsymbol{k}$. Moreover, since $\boldsymbol{e}_{\boldsymbol{k}\lambda}\perp\boldsymbol{k}$, we actually only get a contribution from $\boldsymbol{k}'=\boldsymbol{k}$ since

$$1+\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}'=\begin{cases}2\text{ if }\boldsymbol{k}'=\boldsymbol{k}\\0\text{ if }\boldsymbol{k}'=-\boldsymbol{k}.\end{cases} \tag{8.6}$$

Finally, since $\boldsymbol{e}_{\boldsymbol{k}\lambda}\cdot\boldsymbol{e}_{\boldsymbol{k}\lambda'}=\delta_{\lambda\lambda'}$ (orthogonal polarization vectors), we obtain in total

$$\mathcal{H}=\frac{1}{2}\sum_{\boldsymbol{k}\lambda}\hbar\omega_k(a_{\boldsymbol{k}\lambda}a^*_{\boldsymbol{k}\lambda}+a^*_{\boldsymbol{k}\lambda}a_{\boldsymbol{k}\lambda})=\sum_{\boldsymbol{k}\lambda}\hbar\omega_k a_{\boldsymbol{k}\lambda}a^*_{\boldsymbol{k}\lambda} \tag{8.7}$$

where we used that $a^*_{\boldsymbol{k}\lambda}$ is just a scalar amplitude and thus commutes with $a_{\boldsymbol{k}\lambda}$. It is now useful to introduce the real and canonical variables:

$$q_{\boldsymbol{k}\lambda}=\sqrt{\frac{\hbar}{2\omega_k}}(a_{\boldsymbol{k}\lambda}+a^*_{\boldsymbol{k}\lambda}),\ p_{\boldsymbol{k}\lambda}=\frac{1}{\mathrm{i}}\sqrt{\frac{\hbar\omega_k}{2}}(a_{\boldsymbol{k}\lambda}-a^*_{\boldsymbol{k}\lambda}). \tag{8.8}$$

To prove that these are indeed canonical, recall that $a_{\boldsymbol{k}\lambda}\propto\mathrm{e}^{-\mathrm{i}\omega_k t}$ which provides

$$\begin{aligned}\dot{q}_{\boldsymbol{k}\lambda}&=-\mathrm{i}\sqrt{\frac{\hbar\omega_k}{2}}(a_{\boldsymbol{k}\lambda}-a^*_{\boldsymbol{k}\lambda})=p_{\boldsymbol{k}\lambda}=\frac{\partial\mathcal{H}}{\partial p_{\boldsymbol{k}\lambda}},\\\dot{p}_{\boldsymbol{k}\lambda}&=-\sqrt{\frac{\hbar\omega_k}{2}}(a_{\boldsymbol{k}\lambda}+a^*_{\boldsymbol{k}\lambda})=-\omega_k^2 q_{\boldsymbol{k}\lambda}=-\frac{\partial\mathcal{H}}{\partial q_{\boldsymbol{k}\lambda}},\end{aligned} \tag{8.9}$$

which are precisely Hamilton's equations for canonical variables. To show the last equality in each equation, note that

$$a_{\boldsymbol{k}\lambda}=\frac{1}{\sqrt{2\hbar\omega_k}}(\omega_k q_{\boldsymbol{k}\lambda}+\mathrm{i}p_{\boldsymbol{k}\lambda}),\ a^*_{\boldsymbol{k}\lambda}=\frac{1}{\sqrt{2\hbar\omega_k}}(\omega_k q_{\boldsymbol{k}\lambda}-\mathrm{i}p_{\boldsymbol{k}\lambda}). \tag{8.10}$$

This is a key result: the classical radiation field is formally equivalent to a set of independent harmonic oscillators with mass 1 in our variables.

Quantization.
Having established the equivalence to harmonic oscillators, the procedure to go quantum mechanical is clear: we replace the classical oscillators with QM oscillators. The energy of the system then becomes

$$E = \sum_{\boldsymbol{k}\lambda} (n_{\boldsymbol{k}\lambda} + \frac{1}{2})\hbar\omega_k \tag{8.11}$$

where $n_{\boldsymbol{k}\lambda} = 0, 1, 2, \ldots$ is the number of photons in the mode $(\boldsymbol{k}, \lambda)$. The corresponding state of the system is the product:

$$|n_{\boldsymbol{k}_1,\lambda_1}, n_{\boldsymbol{k}_2,\lambda_2}, \ldots\rangle = |n_{\boldsymbol{k}_1,\lambda_1}\rangle \cdot |n_{\boldsymbol{k}_2,\lambda_2}\rangle \cdot \ldots = \prod_{\boldsymbol{k},\lambda} |n_{\boldsymbol{k},\lambda}\rangle. \tag{8.12}$$

Moreover, we introduce (as for the standard QM harmonic oscillator) creation and annihilation operators $(a^\dagger, a)$ via:

$$q_{\boldsymbol{k}\lambda} = \sqrt{\frac{\hbar}{2\omega_k}}(a_{\boldsymbol{k}\lambda} + a^\dagger_{\boldsymbol{k}\lambda}),\ p_{\boldsymbol{k}\lambda} = \frac{1}{\mathrm{i}}\sqrt{\frac{\hbar\omega_k}{2}}(a_{\boldsymbol{k}\lambda} - a^\dagger_{\boldsymbol{k}\lambda}). \tag{8.13}$$

From the commutator $[p_{\boldsymbol{k}\lambda}, q_{\boldsymbol{k}\lambda}] = \hbar/\mathrm{i}$, it follows that $[a_{\boldsymbol{k}\lambda}, a^\dagger_{\boldsymbol{k}\lambda}] = 1$. The operators for two *different* modes commute, as they are independent of each other. The operators have the known properties (omitting indices for brevity):

$$a^\dagger|n\rangle = \sqrt{n+1}\mathrm{e}^{\mathrm{i}\omega t}|n+1\rangle,\ a|n\rangle = \sqrt{n}\mathrm{e}^{-\mathrm{i}\omega t}|n-1\rangle, \tag{8.14}$$

so that $a^\dagger_{\boldsymbol{k}\lambda}$ creates a photon in the mode $(\boldsymbol{k}, \lambda)$ while $a_{\boldsymbol{k}\lambda}$ removes one such photon. The number operator $N_{\boldsymbol{k}\lambda} = a^\dagger_{\boldsymbol{k}\lambda}a_{\boldsymbol{k}\lambda}$ counts the number of photons in mode $(\boldsymbol{k}, \lambda)$:

$$N_{\boldsymbol{k}\lambda}|\ldots, n_{\boldsymbol{k}\lambda}, \ldots\rangle = n_{\boldsymbol{k}\lambda}|\ldots, n_{\boldsymbol{k}\lambda}, \ldots\rangle. \tag{8.15}$$

The QM operator for the vector potential can now be expressed via creation and annihilation operators:

$$\boldsymbol{A} = \sum_{\boldsymbol{k}\lambda} \boldsymbol{e}_{\boldsymbol{k}\lambda}\sqrt{\frac{\hbar}{2\epsilon_0 V\omega_k}}[a_{\boldsymbol{k}\lambda}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} + a^\dagger_{\boldsymbol{k}\lambda}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}]. \tag{8.16}$$

Calculating the Hamilton-operator in the same way as the classical procedure then yields:

$$\mathcal{H} = \frac{1}{2}\sum_{\boldsymbol{k}\lambda}(a_{\boldsymbol{k}\lambda}a^\dagger_{\boldsymbol{k}\lambda} + a^\dagger_{\boldsymbol{k}\lambda}a_{\boldsymbol{k}\lambda}). \tag{8.17}$$

Note that we cannot any longer freely exchange the order of a and $a^\dagger$, since they are operators that do not commute. Instead, we get:

$$\boxed{\mathcal{H} = E_0 + \sum_{\boldsymbol{k}\lambda}\hbar\omega_k a^\dagger_{\boldsymbol{k}\lambda}a_{\boldsymbol{k}\lambda} \text{ where } E_0 = \frac{1}{2}\sum_{\boldsymbol{k}\lambda}\hbar\omega_k.}$$

Here, E_0 is the ground-state energy (sometimes referred to as the zero-point energy) of the radiation field. Note that the generator for the electric field $\boldsymbol{\mathcal{E}}$ is obtained via $\boldsymbol{\mathcal{E}} = -\partial_t\boldsymbol{A}$. The constant E_0 can *usually* simply be removed since we can choose the reference level for energy where we like. However, there are interesting exceptions such as the Casimir effect. The essence of this phenomenon is that altering the geometry of a system (such as two metallic plates) changes the allowed frequency spectrum $\{\omega_{\boldsymbol{k}\lambda}\}$ and thus changes E_0. If E_0 is reduced, it causes the system to try to alter its geometry, leading *e.g.* to an attraction of the metallic plates.

B. Coherent states

If we compute the expectation value of $\boldsymbol{\mathcal{E}}$ in a state $|n_{\boldsymbol{k}\lambda}\rangle$ for a mode of the radiation field, we obtain

$$\langle n_{\boldsymbol{k}\lambda}|\boldsymbol{\mathcal{E}}|n_{\boldsymbol{k}\lambda}\rangle = 0, \tag{8.18}$$

since $\langle n|a|n\rangle = \langle n|a^\dagger|n\rangle = 0$. That does not seem very encouraging in terms of correctly describing an EM wave. The question becomes: what kind of QM state for the radiation field gives a description which seems more reconcilable with the classical picture?

Our previous treatment of a harmonic oscillator again provides the solution: coherent states $|\alpha\rangle = \sum_{n=0}^{\infty} c_n|n\rangle$ corresponded to a classical oscillation. Due to our established analogy between the QM treatment of the $\boldsymbol{A}$ field and a harmonic oscillator, we can in the same way construct coherent photon states *for each mode* $(\boldsymbol{k}, \lambda)$. Using our previously derived result

$$\langle\alpha|a|\alpha\rangle = \alpha \mathrm{e}^{-\mathrm{i}\omega t},\ \langle\alpha|a^\dagger|\alpha\rangle = a^* \mathrm{e}^{\mathrm{i}\omega t}, \tag{8.19}$$

it follows that the expectation value of the electric field operator $\boldsymbol{\mathcal{E}}$ in a coherent state for the mode $(\boldsymbol{k}, \lambda)$ becomes:

$$\langle\alpha|\boldsymbol{\mathcal{E}}|\alpha\rangle = \mathrm{i}\boldsymbol{e}_{\boldsymbol{k}\lambda}\sqrt{\frac{\hbar\omega_k}{2\epsilon_0 V}}[\alpha \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}-\mathrm{i}\omega_k t} - \alpha^* \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}+\mathrm{i}\omega_k t}]. \tag{8.20}$$

The correspondence to a classical harmonic wave becomes more clear if we write $\alpha = |\alpha|\mathrm{e}^{\mathrm{i}\Theta}$:

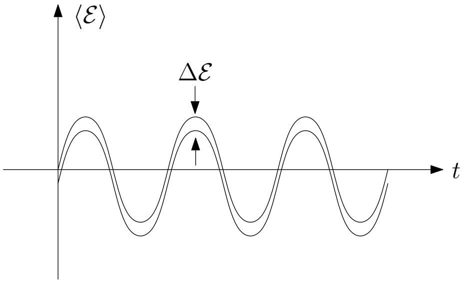
$$\langle\alpha|\boldsymbol{\mathcal{E}}|\alpha\rangle = -\boldsymbol{e}_{\boldsymbol{k}\lambda}\sqrt{\frac{2\hbar\omega_k}{\epsilon_0 V}}|\alpha|\sin(\boldsymbol{k}\cdot\boldsymbol{r} - \omega_k t + \Theta). \tag{8.21}$$

The importance of coherent states lies not only in the fact that they provide a clear, formal similarity between the expectation value of $\boldsymbol{\mathcal{E}}$ in the QM treatment and a classical wave, but also because a monochromatic (fixed wavelength) laser can generate such coherent excitations. Thus, these states have direct experimental relevance.

An interesting observation is that the standard deviation:

$$\Delta\mathcal{E} = \sqrt{\langle \boldsymbol{\mathcal{E}}^2\rangle - \langle \boldsymbol{\mathcal{E}}\rangle^2} \tag{8.22}$$

from the expectation value is *independent* of the field amplitude α. One finds $\Delta\mathcal{E} = \sqrt{\frac{\hbar\omega_k}{2\epsilon_0 V}}$ which thus becomes less and less relevant as $\langle \boldsymbol{\mathcal{E}}\rangle$ increases as shown in the figure.



We can also compute the average number of photons in a coherent state $|\alpha\rangle$ for the mode $(\boldsymbol{k}, \lambda)$:

$$\langle n\rangle = \langle\alpha|N_{\boldsymbol{k}\lambda}|\alpha\rangle = |\alpha|^2. \tag{8.23}$$

The entire distribution of photons [the probability $P(n)$ to find the mode excited with n photons] is found in the usual way: projecting the total state $|\alpha\rangle$ on the state $|n\rangle$ with n photons,

$$P(n) = |\langle n|\alpha\rangle|^2 = \mathrm{e}^{-|\alpha|^2}\frac{|\alpha|^{2n}}{n!}. \tag{8.24}$$

The photon-number in a monochromatic coherent state is thus Poisson-distributed:

$$P(n) = \mathrm{e}^{-\langle n\rangle}\frac{\langle n\rangle^n}{n!}. \tag{8.25}$$

C. Fully quantized radiation theory

We are now in a position to treat both subsystems which are part of radiation theory (atoms and photons) quantum mechanically. We do so using perturbation theory. The unperturbed Hamiltonian H_0 contains no interaction between the two subsystems. We may then write the total state as the product of independent states for each part. It is natural to use the energy states as basis vectors:

$$|\text{atomic + photon system}\rangle = |\text{atomic system}\rangle \cdot |\text{radiation field}\rangle = |\psi\rangle \cdot |n_{\boldsymbol{k}_1,\lambda_1}, n_{\boldsymbol{k}_2,\lambda_2}, \ldots\rangle. \tag{8.26}$$

The two subsystems are coupled via interaction terms such as

$$H_1' = -\frac{q}{m}\boldsymbol{A}\cdot\hat{\boldsymbol{p}} \text{ and } H_3' = -\frac{q}{m}\boldsymbol{S}\cdot(\nabla\times\boldsymbol{A}), \tag{8.27}$$

causing transitions between the unperturbed states. The transition rates can be computed via time-dependent perturbation theory. We now consider some examples.

Spontaneous and stimulated emission.
Spontaneous (stimulated) emission is the emission of light from an excited atom in the absence (presence) of photons in the initial state. Let $(\boldsymbol{k}, \lambda)$ be wavevector and polarization for the emitted photon. The excited state $|\psi_2(t)\rangle$ has energy E_2 and the final state $|\psi_1(t)\rangle$ has energy E_1. Thus, we may write:

$$|i\rangle = |\psi_2\rangle \cdot |\ldots, n_{\boldsymbol{k}\lambda}, \ldots\rangle,\ |f\rangle = |\psi_1\rangle \cdot |\ldots, n_{\boldsymbol{k}\lambda}+1, \ldots\rangle. \tag{8.28}$$

Assume that the dominant perturbation term is:

$$H_1' = \frac{e}{m}\boldsymbol{A}\cdot\hat{\boldsymbol{p}} = \frac{e}{m}\sum_{\boldsymbol{k}\lambda}\sqrt{\frac{\hbar}{2V\epsilon_0\omega_k}}[a_{\boldsymbol{k}\lambda}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} + a_{\boldsymbol{k}\lambda}^\dagger\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}]\boldsymbol{e}_{\boldsymbol{k}\lambda}\cdot\hat{\boldsymbol{p}}. \tag{8.29}$$

Matrix elements involving the photon states are simple since only the term with $a^\dagger_{\boldsymbol{k}\lambda}$ contributes:

$$\langle \ldots, n_{\boldsymbol{k}\lambda}+1, \ldots |H_1'| \ldots, n_{\boldsymbol{k}\lambda}, \ldots\rangle = \frac{e}{m}\sqrt{\frac{\hbar}{2V\epsilon_0\omega_k}}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}+\mathrm{i}\omega_k t}\sqrt{n_{\boldsymbol{k}\lambda}+1}(\boldsymbol{e}_{\boldsymbol{k}\lambda}\cdot\boldsymbol{p}). \tag{8.30}$$

This means that (now including the time dependence from $|\psi_i\rangle$):

$$\langle f|H_1'|i\rangle = \frac{e}{m}\sqrt{\frac{\hbar(n_{\boldsymbol{k}\lambda}+1)}{2V\epsilon_0\omega_k}}\boldsymbol{e}_{\boldsymbol{k}\lambda}\cdot\boldsymbol{M}\mathrm{e}^{-\mathrm{i}(E_2-E_1-\hbar\omega_k)t/\hbar}, \tag{8.31}$$

where $\boldsymbol{M} = \langle\psi_1|\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\boldsymbol{p}|\psi_2\rangle$. Using our results derived previously in chapter 4 for the transition rate between two states, we obtain:

$$W_{2\to1} = \frac{2\pi}{\hbar}\frac{e^2}{m^2}\frac{\hbar}{2V\epsilon_0\omega_k}(n_{\boldsymbol{k}\lambda}+1)|\boldsymbol{e}_{\boldsymbol{k}\lambda}\cdot\boldsymbol{M}|^2\delta(E_2-E_1-\hbar\omega_k). \tag{8.32}$$

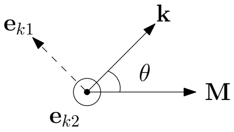
The δ-function ensures energy conservation. We can also obtain the total transition rate from atomic state 2 to atomic state 1, regardless of the mode of the emitted photon, by performing a summation over all possible modes. We use that there are $V/(2\pi)^3 d^3k$ modes with wavevector $\boldsymbol{k}$ in the element d^3k for a given polarization, which yields:

$$\begin{aligned}\omega_{2\to1} &= \sum_\lambda\int W_{2\to1}V\frac{d^3k}{(2\pi)^3}\\ &= \frac{e^2}{m^2}\sum_\lambda\int\frac{d^3k}{(2\pi)^3}\frac{\pi}{\epsilon_0\omega_k}(n_{\boldsymbol{k}\lambda}+1)\delta(E_2-E_1-\hbar\omega_k)|\boldsymbol{e}_{\boldsymbol{k}\lambda}\cdot\boldsymbol{M}|^2.\end{aligned} \tag{8.33}$$

Using that $d^3k = d\boldsymbol{k} = 2\pi\sin\theta d\theta k^2dk$ and $\delta(E_2-E_1-\hbar\omega_k) = \delta(\hbar\omega_{21}-\hbar kc)$ where $\hbar\omega_{21}\equiv E_2-E_1$, we obtain

$$\omega_{2\to1} = \sum_\lambda\frac{e^2\omega_{21}}{4\pi\epsilon_0m^2\hbar c^3}\int_0^\pi\sin\theta d\theta|\boldsymbol{e}_{\boldsymbol{k}\lambda}\cdot\boldsymbol{M}|^2(n_{\boldsymbol{k}\lambda}+1), \tag{8.34}$$

where $k=\omega_{21}/c$. The angle θ specifies the direction of $\boldsymbol{k}$ relative $\boldsymbol{M}$ as shown in the figure.



Spontaneous emission.
Let us for now focus on the case of spontaneous emission $n_{\boldsymbol{k}\lambda}=0$. The summation over polarization directions is easy if we assume that $\boldsymbol{e}_{\boldsymbol{k}1}$ lies in the plane spanned by $\boldsymbol{k}$ and $\boldsymbol{M}$, so that $\boldsymbol{e}_{\boldsymbol{k}2}$ is $\perp$ this plane. Since $\boldsymbol{e}_{\boldsymbol{k}2}\perp\boldsymbol{M}$, we obtain

$$\sum_\lambda|\boldsymbol{e}_{\boldsymbol{k}\lambda}\cdot\boldsymbol{M}|^2 = |\boldsymbol{e}_{\boldsymbol{k}1}\cdot\boldsymbol{M}|^2 = \sin^2\theta|\boldsymbol{M}|^2. \tag{8.35}$$

Performing the resulting angular integration yields as our final result:

$$\omega_{2\to1} = \frac{e^2\omega_{21}}{3\pi\epsilon_0m^2\hbar c^3}|\langle\psi_1|\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\boldsymbol{p}|\psi_2\rangle|^2. \tag{8.36}$$

The name "spontaneous emission" was given during a time when one believed that the process was truly not caused by any interaction. We now see that this point of view is incorrect: the emission occurs as a result of stimulation by the EM field in vacuum.

Stimulated emission.
By setting $n_{\boldsymbol{k}\lambda}=0$ in the factor $(n_{\boldsymbol{k}\lambda}+1)$, we omitted the possibility of stimulated emission: the presence of photons. A key difference from spontaneous emission is that stimulated photons will have the same direction and polarization as the stimulation, rather than being arbitrary. This is a crucial principle behind how a laser works. The opposite process, stimulated absorption, has a transition rate proportional to $n_{\boldsymbol{k}\lambda}$ since:

$$|\langle n_{\boldsymbol{k}\lambda}-1|a_{\boldsymbol{k}\lambda}|n_{\boldsymbol{k}\lambda}\rangle|^2 = n_{\boldsymbol{k}\lambda}. \tag{8.37}$$

IX. DENSITY MATRIX AND QUANTUM STATISTICS

The lecture notes forming the basis for this chapter follow roughly the same structure as the beginning of the corresponding chapters in "Quantum Mechanics" by Bransden & Joachain.

Learning goals. After reading this chapter, the student should:

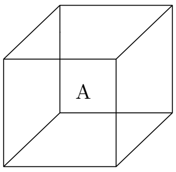
- Be able to describe the difference between mixed states and pure states (in particular a superposition of states).
- Know how to define the density matrix, density operator, and how these quantities can be used to tell apart mixed states from pure states and compute expectation values of operators.
- Be able to qualitatively sketch how information about the polarization of a spin-1/2 system can be obtained using the density matrix.

So far, we have considered quantum systems described by a single wavefunction (or state vector). Such systems are said to be in a *pure state*. These have been assumed prepared in a specific way so that the state vector is completely known. We will now study quantum systems that have states which are incompletely known: *mixed states*. Instead of a single wavefunction, one must use a statistical mixture of wavefunctions to describe such systems.

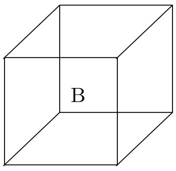
Thus: quantum statistics dealing with such quantum systems as mixed states is the quantum analogue of classical statistical mechanics. A crucial fact that must be strongly emphasized is:

A mixed state is *not the same* as a superposition of states.

Let us illustrate this with a concrete example. Two identical boxes A and B contain a large number of spin-1/2 particles.



100% of the particles are
in the state $\frac{1}{\sqrt{2}}(|+\rangle_z + |-\rangle_z)$



50% of the particles are in state $|+\rangle_z$
50% of the particles are in state $|-\rangle_z$

Which statement is then true?

1. The boxes are the same: the difference is just semantics.
2. The boxes are technically different, but experimentally indistinguishable.
3. The boxes are experimentally different.

Take a minute to think about this. The correct answer is 3, since A is in a pure state (superposition of states) while B is in a mixed state. We can prove this as follows.

Consider a so-called Stern-Gerlach (S-G) device which effectively measures spin in a given direction. If we use a S-G device oriented in the z-direction, A and B give identical results. However, if the S-G device is oriented in the x-direction (so that it measures spin polarized along the x-axis), all particles in box A are measured to be spin-up whereas approximately half the particles in box B are measured to be spin-up. The other half in box B is measured to be spin-down. In effect, boxes A and B give experimentally different results. This can be understood by noting that $|+\rangle_z + |-\rangle_z$ is the $+\hbar/2$ eigenstate of $\hat{S}_x$ while $|+\rangle_z$ and $|-\rangle_z$ individually may be written as 50-50 linear combinations of the $\pm\hbar/2$ eigenstates of $\hat{S}_x$, in effect:

$$|+\rangle_z = \frac{(|+\rangle_z + |-\rangle_z) + (|+\rangle_z - |-\rangle_z)}{2}. \tag{9.1}$$

In order to be able to distinguish clearly mathematically between pure states (which can be superpositions) and mixed states, we will begin by introducing the density matrix formalism. As a concrete application, we shall analyze spin-1/2 particles.

A. The density matrix

Consider a system consisting of an ensemble (collection) of N sub-systems $\alpha = 1, 2, \ldots N$. Suppose that each sub-system is described by a pure state $\psi^{(\alpha)}$. Using Dirac notation, we denote this pure state by $|\alpha\rangle$. All state vectors are assumed normalized to unity, but need not be orthogonal to each other.

Next, select a complete set of basis vectors $|n\rangle$, i.e. orthonormal eigenvectors of some complete set of operators. We then know that $\langle n'|n\rangle = \delta_{n'n}$ and $\sum_n |n\rangle\langle n| = 1$. We expand the pure state $|\alpha\rangle$ in these basis states $|n\rangle$:

$$|\alpha\rangle = \sum_n c_n^{(\alpha)}|n\rangle \to c_n^{(\alpha)} = \langle n|\alpha\rangle. \tag{9.2}$$

Moreover, since $\langle\alpha|\alpha\rangle = 1 \to \sum_n |c_n^{(\alpha)}|^2 = 1$. Consider an observable represented by an operator A. The expectation value in state $|\alpha\rangle$:

$$\langle A\rangle_\alpha = \langle\alpha|A|\alpha\rangle = \sum_{nn'}\langle n|\alpha\rangle\langle\alpha|n'\rangle\langle n'|A|n\rangle. \tag{9.3}$$

Now, the average value of A in the ensemble of pure states is called the *ensemble (or statistical) average* of A and is given by:

$$\langle A\rangle = \sum_{\alpha=1}^{N} W_\alpha\langle A\rangle_\alpha \tag{9.4}$$

where W_α is the statistical weight of each pure state $|\alpha\rangle$, i.e. the probability of finding the system in this state $(0 \le W_\alpha \le 1)$. Clearly, $\sum_{\alpha=1}^{N} W_\alpha = 1$. We then have:

$$\langle A\rangle = \sum_{\alpha=1}^{N}\sum_{nn'}\langle n|\alpha\rangle W_\alpha\langle\alpha|n'\rangle\langle n'|A|n\rangle. \tag{9.5}$$

Let us now introduce the *density operator* $\hat{\rho}$:

$$\hat{\rho} = \sum_{\alpha=1}^{N}|\alpha\rangle W_\alpha\langle\alpha|. \tag{9.6}$$

Taking matrix elements of the density operator between basis states $|n\rangle$, we obtain the density matrix ρ in the $\{n\}$ representation whose elements are:

$$\begin{aligned}\rho_{nn'} &\equiv \langle n|\hat{\rho}|n'\rangle \\ &= \sum_{\alpha=1}^{N}\langle n|\alpha\rangle W_\alpha\langle\alpha|n'\rangle = \sum_{\alpha=1}^{N} W_\alpha[c_{n'}^{(\alpha)}]^* c_n^{(\alpha)}.\end{aligned} \tag{9.7}$$

We emphasize that we are denoting the density *operator* by $\hat{\rho}$ while the density matrix is ρ. Note that the density operator is independent of the choice of the representation, but the density matrix has a different form in different representations. We can thus express $\langle A\rangle$ as follows:

$$\begin{aligned}\langle A\rangle &= \sum_{\alpha=1}^{N}\sum_{nn'} W_\alpha[c_{n'}^{(\alpha)}]^* c_n^{(\alpha)}\langle n'|A|n\rangle \\ &= \sum_{nn'}\langle n|\hat{\rho}|n'\rangle\langle n'|A|n\rangle \\ &= \sum_n\langle n|\hat{\rho}A|n\rangle \\ &= \mathrm{Tr}(\rho A)\end{aligned} \tag{9.8}$$

according to our definition of the matrix elements Eq. (9.7). We have then found that:

Knowing the density matrix enables us to obtain the ensemble average of a quantity A.

We see that a normalization condition $\mathrm{Tr}(\rho) = 1$ is obtained by setting $A = 1$ (identity operator). If we had pure states $|\alpha\rangle$ that were not normalized to unity, then the calculation would have given:

$$\langle A\rangle = \frac{\mathrm{Tr}(\rho A)}{\mathrm{Tr}(\rho)}. \tag{9.9}$$

The density matrix is Hermitian, as seen from its definition: $\langle n|\hat{\rho}|n'\rangle = \langle n'|\hat{\rho}|n\rangle^*$.

A consequence of this is that we can always diagonalize ρ by means of a unitary transformation. Its diagonal elements $\rho_{nn} = \sum_{\alpha=1}^{N} W_\alpha |c_n^{(\alpha)}|^2$ have a simple physical interpretation. They are the probability of finding a member of the ensemble in the pure state $|n\rangle$. We also see from the equations that $\rho_{nn} \geq 0$: $\hat{\rho}$ is a so-called positive semi-definite operator.

Since $\mathrm{Tr}(\rho) = 1$ and $\rho_{nn} \geq 0$, it follows that $0 \leq \rho_{nn} \leq 1$. Moreover, $\mathrm{Tr}(\rho^2) \leq \mathrm{Tr}(\rho) = 1$ because of this. This relation holds regardless of which representation we write the density matrix in, since Tr is invariant under a unitary transformation due to its cyclic property:

$$\mathrm{Tr}(U\rho U^\dagger) = \mathrm{Tr}(UU^\dagger\rho) = \mathrm{Tr}(\rho). \tag{9.10}$$

Consider the special case such that the system is in a particular pure states $|\lambda\rangle$. Then, $W_\alpha = \delta_{\alpha\lambda}$ and from our definition $\hat{\rho} = \sum_{\alpha=1}^{N} |\alpha\rangle W_\alpha \langle\alpha|$, we have $\hat{\rho}^\lambda \equiv \hat{\rho} = |\lambda\rangle\langle\lambda|$. This is called a projection operator onto the state $|\lambda\rangle$ which satisfies $(\rho^\lambda)^2 = \rho^\lambda \to \mathrm{Tr}[(\rho^\lambda)^2] = \mathrm{Tr}(\rho^\lambda) = 1$.

The equation $\mathrm{Tr}[(\rho^\lambda)^2] = 1$ in fact gives us a criterion for *deciding whether a state is pure or not*, and this criterion is invariant under all unitary transformations since Tr is invariant under these. It also follows that

$$\mathrm{Tr}(\rho^\lambda A) = \sum_{nn'} \langle n|\hat{\rho}^\lambda|n'\rangle\langle n'|A|n\rangle = \langle\lambda|A|\lambda\rangle. \tag{9.11}$$

Suppose that we use a representation $\{|k\rangle\}$ in which ρ^λ is diagonal. Then, the above equation is satisfied if:

$$\rho^\lambda_{kk} = \delta_{k\lambda}. \tag{9.12}$$

Then, ρ^λ only has one non-vanishing matrix element which is equal to 1 in the λ-th row and column. In turn, this means that all eigenvalues of the pure state density matrix ρ^λ are equal to zero in *any representation* except one eigenvalue which is equal to unity (since eigenvalues don't change under unitary transformations). This is an equivalent way of characterizing a pure state via the density matrix.

Upon labelling the rows and columns of ρ with indices n and n', both n and n' generally refer to a *set* of indices (such as quantum numbers). Often times, however, we are only interested in a particular property of the system, such as its spin. We then omit the dependence of ρ on all other variables, keeping only the relevant spin variables and in this manner define a *reduced density matrix*.

B. Spin 1/2 system density matrices and polarization

We shall now apply the general methods presented above on the case of spin-1/2 particles, *e.g.* a beam of electrons. The pure states of a spin-1/2 particle are labelled by momentum eigenvalues (p_x, p_y, p_z) and the spin projection eigenvalues $m_s\hbar$ with $m_s = \pm 1/2$. Let $\boldsymbol{z}$ be the quantization axis. The states are then $|p_x, p_y, p_z, m_s\rangle$ and the density matrix elements are:

$$\langle n|\hat{\rho}|n'\rangle = \langle p_x, p_y, p_z, m_s|\hat{\rho}|p'_x, p'_y, p'_z, m'_s\rangle. \tag{9.13}$$

The momentum indices are continuous whereas the spin indices are discrete. We focus here on the spin properties - disregard the momentum labels and look at the reduced density matrix $\langle m_s|\hat{\rho}|m'_s\rangle$, which then is a 2×2 matrix in spin space.

Consider two beams of electrons. One beam has N_a electrons in the pure state $|\chi^a\rangle$. The other beam has N_b electrons in the pure state $|\chi^b\rangle$. The density operator describing the joint beam is:

$$\hat{\rho} = W_\alpha|\chi^a\rangle\langle\chi^a| + W_b|\chi^b\rangle\langle\chi^b| \tag{9.14}$$

where the statistical weights are

$$W_a = \frac{N_a}{N_a + N_b},\ W_b = \frac{N_b}{N_a + N_b}. \tag{9.15}$$

We now choose a basis set of two states $|\chi_1\rangle$ and $|\chi_2\rangle$, for instance the two basic spinors:

$$|\chi_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ |\chi_2\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{9.16}$$

and expand our pure states in terms of these:

$$|\chi^a\rangle = c^a_1|\chi_1\rangle + c^a_2|\chi_2\rangle,\ |\chi^b\rangle = c^b_1|\chi_1\rangle + c^b_2|\chi_2\rangle. \tag{9.17}$$

It follows that the density matrix in the $\{|\chi_i\rangle\}$ representation is given by

$$\rho = \begin{bmatrix} W_a|c^a_1|^2 + W_b|c^b_1|^2 & W_a c^a_1(c^a_2)^* + W_b c^b_1(c^b_2)^* \\ W_a(c^a_1)^* c^a_2 + W_b(c^b_1)^* c^b_2 & W_a|c^a_2|^2 + W_b|c^b_2|^2 \end{bmatrix}. \tag{9.18}$$

If our mixture consisted of N_1 electrons in the $|\chi^a\rangle = |\chi_1\rangle$ state and N_2 electrons in the $|\chi^b\rangle = |\chi_2\rangle$ state, the joint beam would be represented by the density operator

$$\hat{\rho} = W_1|\chi_1\rangle\langle\chi_1| + W_2|\chi_2\rangle\langle\chi_2| \tag{9.19}$$

where $W_1 = N_1/(N_1 + N_2)$ and $W_2 = N_2/(N_1 + N_2)$. Since now $c^a_1 = c^b_2 = 1$ and $c^b_1 = c^a_2 = 0$, the density matrix becomes diagonal in the $\{|\chi_i\rangle\}$ representation:

$$\rho = \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix}. \tag{9.20}$$

Polarization.
Let ρ be a general 2×2 density matrix describing a spin-1/2 system. The unit matrix and three Pauli matrices form a complete set of 2×2 matrices, so we may write in general

$$\rho = a_0 I + a_x \sigma_x + a_y \sigma_y + a_z \sigma_y = a_0 I + \boldsymbol{a} \cdot \boldsymbol{\sigma}. \tag{9.21}$$

Here, $\{a_0, a_x, a_y, a_z\}$ are presently unknown parameters and I is the identity matrix. Since we know that $\text{Tr}(\rho) = 1$ is always satisfied, it follows that $a_0 = 1/2$ by using TrI=2 and Tr$\sigma_i = 0$.

The coefficients a_i, $i = x, y, z$ give information about the polarization of the mixture of states described by ρ. To see this, first note that $\langle \sigma_i \rangle = \text{Tr}(\rho \sigma_i)$ as we have previously derived. Inserting our general expression for ρ and using that $\text{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}$, we find $\langle \sigma_i \rangle = 2a_i$. We can then write:

$$\rho = \frac{1}{2}(I + \boldsymbol{\sigma} \cdot \boldsymbol{P}) \tag{9.22}$$

where $\boldsymbol{P} = \langle \vec{\sigma} \rangle$ is the *polarization vector*. Since ρ is Hermitian, we may always diagonalize it by choosing an appropriate set of basis states. We have

$$\rho = \frac{1}{2}\begin{bmatrix} 1 + P_z & P_x - \mathrm{i}P_y \\ P_x + \mathrm{i}P_y & 1 - P_z \end{bmatrix} \to \rho_{\text{diag}} = \frac{1}{2}\begin{bmatrix} 1 + P & 0 \\ 0 & 1 - P \end{bmatrix}, \tag{9.23}$$

where $P \equiv \pm|\boldsymbol{P}| = \pm\sqrt{P_x^2 + P_y^2 + P_z^2}$. We see that in the representation where ρ is diagonal, one has $P_x = P_y = 0$ and $P = P_z$. Thus, if we let $|\uparrow\rangle$ and $|\downarrow\rangle$ correspond to the kets for spin-up and spin-down with $\boldsymbol{P}$ along the z-axis, we obtain:

$$\boldsymbol{\sigma} \cdot \boldsymbol{P}|\uparrow\rangle = P|\uparrow\rangle,\ \boldsymbol{\sigma} \cdot \boldsymbol{P}|\downarrow\rangle = -P|\downarrow\rangle. \tag{9.24}$$

We previously established the physical interpretation of the elements ρ_{nn}: the probability of finding a member of the ensemble in the pure state $|n\rangle$. It follows in our case that $(1 + P)/2$ is the probability of finding in our mixture the pure states with spin-up along $\boldsymbol{P}$.

This probability may also be expressed as $N_+/(N_+ + N_-)$ where $N_\pm$ is the number of spin measurements giving the value $\pm\hbar/2$ in the $\boldsymbol{P}$-direction. It follows that

$$\frac{1}{2}(1 \pm P) = \frac{N_\pm}{N_+ + N_-} \to P = \frac{N_+ - N_-}{N_+ + N_-}. \tag{9.25}$$

In effect, we have proven that the polarization P is quite naturally the probability of finding the system in the state $|\uparrow\rangle$ minus the probability of finding the system in the state $|\downarrow\rangle$.

If $P = 0$, then $\rho = \text{diag}(1/2, 1/2)$ and the system is in a completely unpolarized and random state. In contrast, we have previously shown that if $\rho^2 = \rho$, the system is in a pure state. When is this the case? We see that:

$$\rho^2 = [\frac{1}{2}(I + \boldsymbol{\sigma} \cdot \boldsymbol{P})]^2 = \frac{1}{4}(I + 2\boldsymbol{\sigma} \cdot \boldsymbol{P} + P^2). \tag{9.26}$$

This is equal to $\frac{1}{2}(I + \boldsymbol{\sigma} \cdot \boldsymbol{P})$ if $P^2 = 1$, which means that there are two pure states corresponding to $P = +1$ and $P = -1$. The physical interpretation is clear: the system is then totally polarized in the direction of $\boldsymbol{P}$ $(P = +1)$ or oppositely to $\boldsymbol{P}$ $(P = -1)$. The corresponding density matrices for pure states with spin projection $\pm\hbar/2$ along $\boldsymbol{z}$ is:

$$\rho = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ for } P = +1 \text{ and } \rho = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ for } P = -1. \tag{9.27}$$

For intermediate values, $0 < |P| < 1$, the system is partially polarized. We conclude this analysis of spin-1/2 systems by commenting on the number of parameters required to determine the density matrix. From our parametrization $\rho = \frac{1}{2}(I + \boldsymbol{\sigma} \cdot \boldsymbol{P})$, it is clear that the 2×2 density matrix for a spin-1/2 mixed state is entirely specified by three real independent parameters $\boldsymbol{P} = (P_x, P_y, P_z)$. In effect, three independent measurements are required to determine ρ for a spin-1/2 system. On the other hand, in the special case of pure states, our previous discussion showed that $P^2 = 1$ is satisfied. This means that only *two* real independent parameters are required in that case since the third is fixed by the condition $P^2 = 1$.