# Notes on Quantum Mechanics 

Finn Ravndal<br>Institute of Physics<br>University of Oslo, Norway<br>e-mail: finnr@fys.uio.no

v.4: December, 2006

## Contents

1 Linear vector spaces ..... 7
1.1 Real vector spaces ..... 7
1.2 Rotations ..... 9
1.3 Lie groups ..... 11
1.4 Change of basis ..... 12
1.5 Complex vector spaces ..... 12
1.6 Unitary transformations ..... 14
2 Basic quantum mechanics ..... 17
2.1 Bra and ket vectors ..... 17
2.2 Operators in Hilbert space ..... 19
2.3 Matrix representations ..... 20
2.4 Adjoint operations ..... 21
2.5 Eigenvectors and eigenvalues ..... 22
2.6 The quantum postulates ..... 23
2.7 Ehrenfest's theorem ..... 25
2.8 The time evolution operator ..... 25
2.9 Stationary states and energy basis ..... 27
2.10 Change of basis ..... 28
2.11 Schrödinger and Heisenberg pictures ..... 29
2.12 The Lie formula ..... 31
2.13 Weyl and Cambell-Baker-Hausdorff formulas ..... 32
3 Discrete quantum systems ..... 35
3.1 Matrix mechanics ..... 35
3.2 The hydrogen molecule ..... 36
3.3 The general two-state system ..... 38
3.4 Neutrino oscillations ..... 40
3.5 Reflection invariance and symmetry transformations ..... 41
3.6 Lattice translation invariance ..... 43
3.7 Lattice dynamics and energy bands ..... 45
3.8 Three-dimensional lattices ..... 47
4 Schrödinger wave mechanics ..... 51
4.1 Coordinate basis ..... 51
4.2 The Dirac $\delta$-function ..... 52
4.3 Continuous translations and momentum ..... 53
4.4 Operators in the coordinate representation ..... 55
4.5 The momentum representation ..... 58
4.6 Three dimensions ..... 58
4.7 The wave equation ..... 60
4.8 Algebraic solution of the harmonic oscillator ..... 61
4.9 Hermite polynomials ..... 66
4.10 Coherent states ..... 67
5 Rotations and spherical symmetry ..... 71
5.1 Cyclic molecules ..... 71
5.2 Particle on a circle ..... 73
5.3 Axial symmetry ..... 74
5.4 Planar harmonic oscillator ..... 76
5.5 Spherical rotations and the generators ..... 79
5.6 Quantization of spin ..... 83
5.7 Spin-1/2 and the Pauli matrices ..... 86
5.8 Orbital angular momentum and spherical harmonics ..... 89
5.9 Spherical symmetric potentials ..... 93
5.10 The spherical well ..... 96
5.11 Hydrogen atoms ..... 97
5.12 Rotation matrices and Euler angles ..... 101
6 Approximation schemes ..... 109
6.1 The variational method ..... 109
6.2 Rayleigh-Ritz scheme ..... 112
6.3 Static perturbations ..... 113
6.4 Degeneracy ..... 116
6.5 Time-dependent perturbations ..... 118
6.6 Harmonic perturbations ..... 121
6.7 Rutherford scattering ..... 122
6.8 Quantum potential scattering ..... 125
7 Electromagnetic interactions ..... 129
7.1 Magnetic moments ..... 129
7.2 Vector potentials and gauge invariance ..... 130
7.3 Gauge transformations in quantum theory ..... 133
7.4 The Pauli equation ..... 134
7.5 Zeeman effect ..... 137
7.6 Absorption and emission of light ..... 138
7.7 Einstein A and B coefficients ..... 142
7.8 Quantization of the electromagnetic field ..... 144
7.9 Spontaneous emission ..... 151
7.10 Angular distributions ..... 154
7.11 Magnetic transitions ..... 156
7.12 Photo-electric effect ..... 158
7.13 Thomson scattering ..... 160
7.14 Vaccum energies and the cosmological constant ..... 163
7.15 The Casimir force ..... 165
8 Spin interactions ..... 169
8.1 Relativistic velocity correction ..... 169
8.2 Spin-orbit splitting ..... 171
8.3 Addition of spins ..... 174
8.4 Fine structure ..... 177
8.5 Anomalous Zeeman effect and Landé g-factor ..... 179
8.6 Hyperfine splitting ..... 182
8.7 Magnetic transitions in hyperfine doublet ..... 186
9 Many-electron atoms ..... 189
9.1 The helium atom ..... 189
9.2 Electron spin and the Pauli principle ..... 193
9.3 Central-field approximation ..... 196
9.4 Terms and the Hund rules ..... 200
9.5 Atomic magnetic moments ..... 204
9.6 Hartree approximations ..... 205
10 Non-relativistic quantum fields ..... 207
10.1 Identical particles ..... 207
10.2 Fock space ..... 210
10.3 Quantum field operators ..... 213
10.4 Inter-particle interactions ..... 217
10.5 Fermions ..... 220
10.6 Particles and holes ..... 222
10.7 Lattice field theory ..... 224

## Chapter 1

## Linear vector spaces

The most general formulation of quantum mechanics can be made within the framework of a linear vector space. This is a generalization of ideas we have about ordinary vectors in three-dimensional Euclidean space. We will need vector spaces of higher dimensions. In fact, ordinary quantum-mechanical wavefunctions will be found to be the components of a state vector in a vector space with infinite dimensions. Another generalization will be to consider vectors with components that are not real, but complex numbers. A new scalar product must then be defined. In order to discuss the components, we need to have a basis in the vector space. The values of the components depend on this choice of basis vectors and are related by linear transformations to the components in another basis. When such transformations conserve the lengths of vectors, they are said to be rotations in a real vector space and unitary transformations in a complex vector space.

### 1.1 Real vector spaces

Let us consider a set $\mathcal{V}$ of vectors $\{\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots\}$ which has the basic property that that sum of any two vectors is a new vector in the same set,

$$
\begin{equation*}
\mathbf{x}+\mathbf{y} \in \mathcal{V} \tag{1.1}
\end{equation*}
$$

Summation is assumed to be commutative, $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$ and associative, $(\mathbf{x}+\mathbf{y})+\mathbf{z}=$ $\mathbf{x}+(\mathbf{y}+\mathbf{z})$. There must also be is null vector $\mathbf{0}$ such that for any vector $\mathbf{x}$ we will have $\mathbf{x}+\mathbf{0}=\mathbf{x}$. The vector which adds to $\mathbf{x}$ to give the null vector, is called the inverse vector. These are the basic definitions of a linear vector space.
Now to get any further, we must assume that we also have a set of scalars $\{a, b, c, \ldots\}$ which we first take to be ordinary, real numbers. Multiplying a vector with such a scalar, we find a new vector, i.e. $a \mathbf{x} \in \mathcal{V}$. The linear combination $a \mathbf{x}+b \mathbf{y}$ is therefore also a vector. Multiplying with zero we get the zero vector. Thus we can write the vector inverse to $\mathbf{x}$ to be $-\mathbf{x}$.

A vector $\mathbf{v}$ is said to be linearly independent of the vectors $\{\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots\}$ when it cannot be written as a linear combination of these vectors. The maximum number of linearly independent vectors is called the dimension $N$ of the vector space which now can be
denoted as $\mathcal{V}_{N}$. Such a set is called a basis for the vector space. If this basis set is denoted by $\left\{\mathbf{e}_{1}, \mathbf{e}_{2}, \ldots, \mathbf{e}_{N}\right\}$, any vector $\mathbf{x}$ can therefore be written as the linear combination $\mathbf{x}=x_{1} \mathbf{e}_{1}+x_{2} \mathbf{e}_{2}+\ldots+x_{N} \mathbf{e}_{N}$ or

$$
\begin{equation*}
\mathbf{x}=\sum_{n=1}^{N} x_{n} \mathbf{e}_{n} \tag{1.2}
\end{equation*}
$$

The scalar coefficients $x_{n}$ are called the components of the vector with respect to this basis. It therefore follows that the components of the $\operatorname{sum} \mathbf{x}+\mathbf{y}$ of two vectors are just the sum $x_{n}+y_{n}$ of their components. Obviously, the components of the null vector are all zero.

We have now constructed a linear vector space over real numbers. Measurements in physics are represented by real numbers. For a mathematical construction like a vector space to be physically useful, we must therefore be able to assign sizes or lengths to all vectors in the space. This can be done by endowing the vector space with an inner or scalar product. In analogy with standard vector analysis, we denote the scalar product of a vector $\mathbf{x}$ with itself by $\mathbf{x} \cdot \mathbf{x}$ which is then called its squared length and is a positive number. Similarly, the scalar product of vector $\mathbf{x}$ with vector $\mathbf{y}$ is written as $\mathbf{x} \cdot \mathbf{y}$. When this product is zero, we say that the two vectors are normal or orthogonal to each other.

All scalar products are given in terms of the scalar products between the vectors in the basis set. It is most convenient to choose a very special set of such vectors where each is orthogonal to all the others and all have the same length which we take to be one. They are therefore unit vectors. In this orthonormal basis we thus have

$$
\mathbf{e}_{m} \cdot \mathbf{e}_{n}=\delta_{m n}= \begin{cases}1, & m=n  \tag{1.3}\\ 0, & m \neq n\end{cases}
$$

where we one the right have introduced the Kroenecker $\delta$-symbol. The scalar product of two vectors $\mathbf{x}$ and $\mathbf{y}$ is then

$$
\begin{equation*}
\mathbf{x} \cdot \mathbf{y}=\sum_{m, n=1}^{N} x_{m} y_{n} \mathbf{e}_{m} \cdot \mathbf{e}_{n}=\sum_{n=1}^{N} x_{n} y_{n} \tag{1.4}
\end{equation*}
$$

and similarly for the squared length of the vector $\mathbf{x}$,

$$
\begin{equation*}
|\mathbf{x}|^{2}=\mathbf{x} \cdot \mathbf{x}=\sum_{n=1}^{N} x_{n}^{2} \tag{1.5}
\end{equation*}
$$

The length of the null vector is therefore null.
Multiplying the vector $\mathbf{x}$ with the basis vector $\mathbf{e}_{n}$, we find the component of the vector in this direction,

$$
\begin{equation*}
x_{n}=\mathbf{e}_{n} \cdot \mathbf{x} \tag{1.6}
\end{equation*}
$$

Inserting this back into the vector (1.2), we have

$$
\mathbf{x}=\sum_{n=1}^{N} \mathbf{e}_{n} \mathbf{e}_{n} \cdot \mathbf{x}
$$

Since this is valid for all possible vectors $\mathbf{x}$, we must have that the object

$$
\begin{equation*}
I=\sum_{n=1}^{N} \mathbf{e}_{n} \mathbf{e}_{n} \tag{1.7}
\end{equation*}
$$

acts as a unit operator or identity in the sense that when it acts on any vector, it just gives the same vector back. In fact, we have just shown that $I \cdot \mathbf{x}=\mathbf{x}$ and obviously also $I \cdot \mathbf{e}_{m}=\mathbf{e}_{m}$ for any basis vector. That the sum in (1.7) is just the identity, is usually said to show that the basis vectors satisfy the completeness relation.

The unit operator $I$ is just one special operator on this vector space. A general operator $S$ acting on a vector $\mathbf{x}$ gives a new vector $\mathbf{x}^{\prime}$, i.e. $\mathbf{x}^{\prime}=S \cdot \mathbf{x}$. We will only consider linear operators defined by $S \cdot(\mathbf{x}+\mathbf{y})=S \cdot \mathbf{x}+S \cdot \mathbf{y}$. It is trivial to see from the definition (1.7) that $I$ is linear. Each operator $S$ will have an inverse denoted by $S^{-1}$ such that $S \cdot S^{-1}=S^{-1} \cdot S=I$.

### 1.2 Rotations

A rotation of a vector $\mathbf{x}$ is a linear transformation of its components so that its length stays the same. The transformed vector $\mathbf{x}^{\prime}=\sum_{n=1}^{N} x_{n}^{\prime} \mathbf{e}_{n}$ will thus have components


Figure 1.1: Rotation by an angle $\alpha$ of a two-dimensional vector $\mathbf{x}$ into the final vector $\mathbf{x}^{\prime}$.
which in general can be written as

$$
\begin{equation*}
x_{m}^{\prime}=\sum_{n=1}^{N} R_{m n} x_{n} \tag{1.8}
\end{equation*}
$$

where the coefficients $R_{m n}$ are real numbers. This is illustrated in Fig. 1 for the simplest case when the vector space has only $N=2$ dimensions. Since we require that $\left|\mathbf{x}^{\prime}\right|=|\mathbf{x}|$
it follows from (1.5) that we must have

$$
\left|\mathbf{x}^{\prime}\right|^{2}=\sum_{m=1}^{N} x_{m}^{\prime} x_{m}^{\prime}=\sum_{m=1}^{N} \sum_{n, n^{\prime}=1}^{N} R_{m n} R_{m n^{\prime}} x_{n} x_{n^{\prime}}
$$

which should equal $|\mathbf{x}|^{2}=\sum_{n=1}^{N} x_{n} x_{n}$. For this to be the case we must therefore have

$$
\begin{equation*}
\sum_{m=1}^{N} R_{m n} R_{m n^{\prime}}=\delta_{n n^{\prime}} \tag{1.9}
\end{equation*}
$$

These relations impose $N(N+1) / 2$ constraints on the $N^{2}$ rotation coefficients $R_{m n}$ so that a general rotation involves only $N(N-1) / 2$ independent parameters.

It's very useful to write the transformation equation (1.8) on matrix form. We then collect the components of the vector $\mathbf{x}$ in the column matrix

$$
x=\left(\begin{array}{c}
x_{1}  \tag{1.10}\\
x_{2} \\
\vdots \\
x_{N}
\end{array}\right)
$$

and all the rotation coefficients in the matrix

$$
R=\left(\begin{array}{cccc}
R_{11} & R_{12} & \ldots & R_{1 N}  \tag{1.11}\\
R_{21} & R_{22} & \ldots & R_{2 N} \\
\vdots & \vdots & \ddots & \vdots \\
R_{N 1} & R_{N 2} & \ldots & R_{N N}
\end{array}\right)
$$

Using the rules for standard matrix multiplication we can therefore write the transformation (1.8) on the compact form

$$
\begin{equation*}
x^{\prime}=R x \tag{1.12}
\end{equation*}
$$

where $R$ is a $N \times N$ real rotation matrix. We say that it represents the corresponding rotation operator. Similarly, introducing the transposed matrix $R^{T}$ with elements $R_{n m}^{T}=$ $R_{m n}$, we can write the constraints (1.9) as

$$
\begin{equation*}
R^{T} R=I \tag{1.13}
\end{equation*}
$$

where $I$ is the $N \times N$ unit matrix with elements $I_{m n}=\delta_{m n}$, i.e. with ones along the diagonal and zeros elsewhere.
If the rotation $R$ is followed by another rotation $R^{\prime}$, the resulting vector is $x^{\prime \prime}=R^{\prime} x^{\prime}=$ $R^{\prime} R x$. This should also be obtainable by rotating the first vector $x$ directly into $x^{\prime \prime}$, i.e. we must have $x^{\prime \prime}=R^{\prime \prime} x$. It therefore follows that we must have $R^{\prime \prime}=R^{\prime} R$. A rotation $R$ followed by the inverse rotation, which we denote by $R^{-1}$, is equivalent to no rotation at all, i.e. should be equal to the unit matrix. We thus have $R^{-1} R=I$. For the same reason we must also have $R R^{-1}=I$. From (1.13) we now see that the inverse rotation matrix is just the transposed,

$$
\begin{equation*}
R^{-1}=R^{T} \tag{1.14}
\end{equation*}
$$

Taking the determinant of the matrix equation (1.13) and remembering that the determinant of the transposed matrix equals the determinant of the matrix itself, we see that we must have $\operatorname{det}(R) \operatorname{det}(R)=1$. Since the unit matrix has determinant $\operatorname{det}(I)=1$, rotation matrices continuously connected to it must therefore also have $\operatorname{det}(R)=1$.

### 1.3 Lie groups

With these properties we say that the rotations form a continuous group or a Lie group after the Norwegian mathematician Sophus Lie. A general group is a set of elements $\left\{e, g, g^{\prime}, g^{\prime \prime}, \ldots\right\}$ which can be combined together ( $=$ multiplied) in such a way that

- the product $g^{\prime} g$ of two group elements $g$ and $g^{\prime}$ is also a group element, $g^{\prime \prime}=g^{\prime} g$.
- there is a unit element $e$ so that $e g=g e=g$ for all elements $g$.
- for each element there is also an inverse element $g^{-1}$ so that $g g^{-1}=g^{-1} g=e$.

When the number of elements in the group is finite, it is called a finite group. On the other hand, it is a Lie group when it contains infinite elements continuously connected to each other by variations of a certain set of finite parameters.

We have already been introduced to general rotations. The corresponding transformation matrices were found to have properties which show that they form a Lie group. The usual name for this rotation group of $N$-dimensional vectors is $S O(N)$ where the $O$ stands for orthogonal transformation and the $S$ for the special case where the determinant $\operatorname{det}(R)=1$. It was shown to have $N(N-1) / 2$ independent parameters.

As a simple example, consider again the rotation in Fig.1.1. If the length of the vector is $a$, the initial vector before rotation is represented by the column vector

$$
x=\binom{a \cos \phi}{a \sin \phi}
$$

where $\phi$ is the angle the vector makes with the 1-axis. The transformed vector has the new angle $\phi^{\prime}=\phi+\alpha$ where $\alpha$ is the rotation angle. Since $\cos (\phi+\alpha)=\cos \phi \cos \alpha-$ $\sin \phi \sin \alpha$ and $\sin (\phi+\alpha)=\sin \phi \cos \alpha+\cos \phi \sin \alpha$, we can write

$$
x^{\prime}=\left(\begin{array}{rr}
\cos \alpha & -\sin \alpha  \tag{1.15}\\
\sin \alpha & \cos \alpha
\end{array}\right)\binom{a \cos \phi}{a \sin \phi}
$$

Comparing with (1.12) we thus find the rotation matrix to be

$$
R(\alpha)=\left(\begin{array}{rr}
\cos \alpha & -\sin \alpha  \tag{1.16}\\
\sin \alpha & \cos \alpha
\end{array}\right)
$$

which is an element of the rotation group $S O(2)$. It has only one parameter which is the rotation angle and $\operatorname{det}(R)=\cos ^{2} \alpha+\sin ^{2} \alpha=1$. By direct multiplication we now verify that $R\left(\alpha^{\prime}\right) R(\alpha)=R\left(\alpha^{\prime \prime}\right)$ where the resulting rotation angle is $\alpha^{\prime \prime}=\alpha^{\prime}+\alpha$ as it should be.

Much more interesting is the rotation group in three dimensions, $S O(3)$. It has $3(3-$ 1) $/ 2=3$ parameters corresponding to rotations around the three space directions. We will come back to several aspects of this group later in connection with the quantization of spin without having to use any details of group theory. For rotations in spaces with more than three dimensions, the number of independent parameters is larger than the number of space dimensions. It is a good problem to understand why.

### 1.4 Change of basis

So far we have considered an active transformation where the vector $\mathbf{x}$ is rotated while the basis vectors $\left\{\mathbf{e}_{n}\right\}$ remains fixed. The components of the vectors thus takes on the new values (1.8) after the rotation. Another possibility is to consider a passive transformation where the vector $\mathbf{x}$ remains fixed, but the basis vectors $\left\{\mathbf{e}_{n}\right\}$ are rotated. The components of the vector with respect to this new basis will then also be changed.

If we now consider the special passive rotation which is just the inverse of the previous, active rotation we have just studied, the new components will again the given by (1.12). Since the vector $\mathbf{x}$ remains fixed, we must have

$$
\begin{equation*}
\mathbf{x}=\sum_{n=1}^{N} x_{n} \mathbf{e}_{n}=\sum_{m=1}^{N} x_{m}^{\prime} \mathbf{e}_{m}^{\prime} \tag{1.17}
\end{equation*}
$$

Inserting here the values of $x_{m}^{\prime}$ from (1.8) and comparing the coefficients of $x_{n}$ on both sides, we find that the old basis vectors can be given in terms of the new ones by the linear combinations

$$
\begin{equation*}
\mathbf{e}_{n}=\sum_{m=1}^{N} \mathbf{e}_{m}^{\prime} R_{m n} \tag{1.18}
\end{equation*}
$$

Since the rotation matrix is orthogonal, we can easily invert this set of equations to find the transformed basis vectors expressed in terms of the original ones,

$$
\begin{equation*}
\mathbf{e}_{m}^{\prime}=\sum_{n=1}^{N} R_{m n} \mathbf{e}_{n} \tag{1.19}
\end{equation*}
$$

It now follows that this transformed basis is also orthonormal, i.e. $\mathbf{e}_{m}^{\prime} \cdot \mathbf{e}_{n}^{\prime}=\delta_{m n}$ as it should be. Depending on the basis we choose to use, the same, abstract vector can therefore be represented by different column vectors containing the components in the relevant basis.

### 1.5 Complex vector spaces

In the previous sections we considered a linear vector space over real numbers. We will now derive the properties of a complex vector space. It is most easily constructed from a real vector space $\mathcal{V}_{2 N}$ with $2 N$ dimensions and orthogonal basis $\left\{\mathbf{e}_{n}\right\}$. A general vector $\mathbf{x} \in \mathcal{V}_{2 N}$ can then be written as

$$
\begin{equation*}
\mathbf{x}=x_{1} \mathbf{e}_{1}+x_{2} \mathbf{e}_{2}+\ldots+x_{2 N} \mathbf{e}_{2 N}=\sum_{n=1}^{2 N} x_{n} \mathbf{e}_{n} \tag{1.20}
\end{equation*}
$$

We now introduce the complex components

$$
\begin{equation*}
z_{k}=\sqrt{\frac{1}{2}}\left(x_{2 k-1}+i x_{2 k}\right) \quad k=1,2, \ldots, N \tag{1.21}
\end{equation*}
$$

and their complex conjugates $z_{k}^{*}$ together with the new, complex basis vectors

$$
\begin{equation*}
\boldsymbol{\epsilon}_{k}=\sqrt{\frac{1}{2}}\left(\mathbf{e}_{2 k-1}-i \mathbf{e}_{2 k}\right) \quad k=1,2, \ldots, N \tag{1.22}
\end{equation*}
$$

and their complex conjugates $\boldsymbol{\epsilon}_{k}^{*}$. The above real vector (1.20) can now be expressed in this complex basis as

$$
\begin{align*}
\mathbf{x} & =z_{1} \boldsymbol{\epsilon}_{1}+z_{1}^{*} \boldsymbol{\epsilon}_{1}^{*}+z_{2} \boldsymbol{\epsilon}_{2}+z_{2}^{*} \boldsymbol{\epsilon}_{2}^{*}+\ldots+z_{N} \boldsymbol{\epsilon}_{N}+z_{N}^{*} \boldsymbol{\epsilon}_{N}^{*} \\
& =\sum_{k=1}^{N} z_{k} \boldsymbol{\epsilon}_{k}+\sum_{k=1}^{N} z_{k}^{*} \boldsymbol{\epsilon}_{k}^{*} \tag{1.23}
\end{align*}
$$

The first part corresponds to a complex vector $\mathbf{z}$ with components $z_{k}$ in the basis $\left\{\boldsymbol{\epsilon}_{k}\right\}$,

$$
\begin{equation*}
\mathbf{z}=\sum_{k=1}^{N} z_{k} \boldsymbol{\epsilon}_{k} \tag{1.24}
\end{equation*}
$$

As previously, we easily see that they define a linear vector space which we denote by $\mathcal{H}_{N}$. We thus have $\mathbf{z} \in \mathcal{H}_{N}$. The second part describes the complex conjugate or dual vector $\mathbf{z}^{*} \in \mathcal{H}_{N}^{*}$ where $\mathcal{H}_{N}^{*}$ is the dual vector space spanned by the dual basis $\left\{\boldsymbol{\epsilon}_{k}^{*}\right\}$. We therefore have $\mathbf{x}=\mathbf{z}+\mathbf{z}^{*}$. Equivalently, the original vector space $\mathcal{V}_{2 N}$ is the direct product of $\mathcal{H}_{N}$ and its dual, i.e. $\mathcal{V}_{2 N}=\mathcal{H}_{N} \otimes \mathcal{H}_{N}^{*}$.
Since our derivation is based on a real vector space with an inner product between vectors, we should also be able to define an inner vector product in the complex vector space $\mathcal{H}_{N}$. But we cannot simply take $\boldsymbol{\epsilon}_{i} \cdot \boldsymbol{\epsilon}_{j}$ since these basis vectors have zero norm, $\boldsymbol{\epsilon}_{k} \cdot \boldsymbol{\epsilon}_{k}=0$ as follows directly from the definition (1.22). Instead we must combine a vector $\mathbf{z}$ from $\mathcal{H}_{N}$ with its dual partner $\mathbf{z}^{*}$ from $\mathcal{H}_{N}^{*}$ so that our complex basis vectors have the scalar or inner product

$$
\begin{equation*}
\boldsymbol{\epsilon}_{i}^{*} \cdot \boldsymbol{\epsilon}_{j}=\delta_{i j} \tag{1.25}
\end{equation*}
$$

Although we cannot any longer so easily talk about angles between vectors, we still say that this is an orthonormal basis. Two complex vectors $\mathbf{z}$ and $\mathbf{z}^{\prime}$ then have the scalar product

$$
\begin{equation*}
\mathbf{z}^{*} \cdot \mathbf{z}^{\prime}=\sum_{i, j=1}^{N} z_{i}^{*} z_{j}^{\prime} \boldsymbol{\epsilon}_{i}^{*} \cdot \boldsymbol{\epsilon}_{j}=\sum_{k=1}^{N} z_{k}^{*} z_{k}^{\prime} \tag{1.26}
\end{equation*}
$$

which is now in general a complex number. They are said to be normal or orthogonal to each other when this product is zero. The squared norm or length of the vector (1.24) is therefore

$$
\begin{equation*}
|\mathbf{z}|^{2}=\sum_{k=1}^{N}\left|z_{k}\right|^{2}=\sum_{i=1}^{2 N} x_{i}^{2}=|\mathbf{x}|^{2} \tag{1.27}
\end{equation*}
$$

and this is obviously a positive number.
Given a vector z, it has the components

$$
\begin{equation*}
z_{k}=\boldsymbol{\epsilon}_{k}^{*} \cdot \mathbf{z} \tag{1.28}
\end{equation*}
$$

in the basis $\left\{\boldsymbol{\epsilon}_{k}\right\}$. Inserting this back in the expression (1.24) for the vector, we see that we must have the completeness relation

$$
\begin{equation*}
\sum_{k=1}^{N} \epsilon_{k} \epsilon_{k}^{*}=I \tag{1.29}
\end{equation*}
$$

where $I$ is the identity operator in the sense that $I \cdot \mathbf{z}=\mathbf{z}$. This is equivalent to the similar completeness relation we found in real vector spaces.

### 1.6 Unitary transformations

Having established the $N$-dimensional, complex vector space $\mathcal{H}_{N}$ with the orthogonal basis (1.25), we can now consider a linear transformation of the vector $\mathbf{z}$ in this space which would correspond to a rotation in a real vector space. The components of the transformed vector $\mathbf{z}^{\prime}$ can then in general be written as

$$
\begin{equation*}
z_{i}^{\prime}=\sum_{j=1}^{N} U_{i j} z_{j} \tag{1.30}
\end{equation*}
$$

where the coefficients $U_{i j}$ are complex numbers. In order for the length of the vector to remain constant under this transformation, we find from (1.27) that we have the following constraints

$$
\sum_{i=1}^{N} \sum_{j, k=1}^{N} U_{i j}^{*} U_{i k} z_{j}^{*} z_{k}=\sum_{k=1}^{N} z_{k}^{*} z_{k}
$$

which these coefficients must satisfy. Since they should hold for all vectors $\mathbf{z}$, we must have

$$
\begin{equation*}
\sum_{i=1}^{N} U_{i j}^{*} U_{i k}=\delta_{j k} \tag{1.31}
\end{equation*}
$$

Again we can write these equations on a more compact form by introducing the complex $N \times N$ matrix

$$
U=\left(\begin{array}{cccc}
U_{11} & U_{12} & \ldots & U_{1 N}  \tag{1.32}\\
U_{21} & U_{22} & \ldots & U_{2 N} \\
\vdots & \vdots & \ddots & \vdots \\
U_{N 1} & U_{N 2} & \ldots & U_{N N}
\end{array}\right)
$$

Defining the adjoint or Hermitian conjugate matrix $U^{\dagger}$ with elements

$$
\begin{equation*}
U_{i j}^{\dagger}=U_{j i}^{*} \tag{1.33}
\end{equation*}
$$

we see that the constraints (1.31) can be summed up in the simple matrix relation

$$
\begin{equation*}
U^{\dagger} U=I \tag{1.34}
\end{equation*}
$$

Such matrices are said to be unitary and play the same role in complex vector spaces as orthogonal matrices play in real vector spaces.
Introducing the complex column vector

$$
z=\left(\begin{array}{c}
z_{1}  \tag{1.35}\\
z_{2} \\
\vdots \\
z_{N}
\end{array}\right)
$$

the squared norm of the vector $\mathbf{z}$ is given by $z^{\dagger} z$. The unitary transformation (1.30) can then be written on the matrix form

$$
\begin{equation*}
z^{\prime}=U z \tag{1.36}
\end{equation*}
$$

exactly as the rotation of a real vector in the previous note.
These transformations also form a Lie group which for obvious reasons is called the unitary group denoted by $U(N)$. From (1.34) we see that the tranformation $U$ has an inverse which is just the adjoint matrix,

$$
\begin{equation*}
U^{-1}=U^{\dagger} \tag{1.37}
\end{equation*}
$$

It has $N^{2}$ complex matrix elements, i.e. $2 N^{2}$ real parameters. But they are not all independent, the requirement (1.34) forms $N^{2}$ real constraints. We thus find that the unitary group $U(N)$ has $N^{2}$ independent and real parameters.

From the constraints (1.31) also follows $\operatorname{det}(U)^{*} \operatorname{det}(U)=1$ by taking the determinant on both sides. Except for a phase factor, we can therefore require that the matrices have $\operatorname{det}(U)=1$. This is then one more real constraint and the resulting special unitary group is called $S U(N)$ with $N^{2}-1$ parameters.

As a first example, let us consider the simplest unitary group which is $U(1)$. Its group elements can all be written as

$$
\begin{equation*}
U(\alpha)=e^{i \alpha} \tag{1.38}
\end{equation*}
$$

where the real parameter $\alpha$ takes value between zero and $2 \pi$. The combination of two such elements $U(\alpha)$ and $U\left(\alpha^{\prime}\right)$ gives the new group element

$$
U\left(\alpha^{\prime \prime}\right)=U\left(\alpha^{\prime}\right) U(\alpha)=U\left(\alpha^{\prime}+\alpha\right)
$$

It acts in a complex vector space with only $N=1$ dimension where the vectors have only one component. As an axample, let it be the complex number $z=x+i y$ which also can be written as $z=a e^{i \phi}=a(\cos \phi+i \sin \phi)$ with the angle $\phi$ between zero and $2 \pi$. The unitary transformation (1.36) then becomes simply

$$
z^{\prime}=e^{i \alpha} a e^{i \phi}=a e^{i \phi^{\prime}}
$$

with $\phi^{\prime}=\phi+\alpha$. Writing $z^{\prime}=x^{\prime}+i y^{\prime}$ and comparing real and imaginary parts on boths sides, we see that $U(1)$ is equivalent to the real rotation group $S O(2)$ with elements (1.16). Even if the group properties of $U(1)$ are rather trivial, it plays an important role
in quantum mechanics and what is called gauge transformations of the electromagnetic field.

Mathematically much more interesting and physically just as important, is the unitary group $S U(2)$. It is defined by $2 \times 2$ unitary matrices with unit determinant and have three real parameters. One can write a general group element as

$$
U=\left(\begin{array}{cc}
a & -b  \tag{1.39}\\
b^{*} & a^{*}
\end{array}\right)
$$

where $a$ and $b$ are complex parameters satisfying $|a|^{2}+|b|^{2}=1$. This ensures that $\operatorname{det}(U)=1$. It is then easy to find the inverse transformation

$$
U^{-1}=\left(\begin{array}{cc}
a^{*} & b  \tag{1.40}\\
-b^{*} & a
\end{array}\right)
$$

which is also the adjoint matrix $U^{\dagger}=U^{* T}$.
We will later see that $S U(2)$ is intimately related to the three-dimensional rotation group $S O(3)$ and gives rise to particles with half-integer spins. As a starter we notice that both groups have the same number of independent parameters. In fact, the transformation matrix (1.39) will rotate particles with spin $S=1 / 2$.

## Chapter 2

## Basic quantum mechanics

In quantum mechanics all the information about the physical state of a system is contained in a state vector. This is a vector in a complex vector space called the Hilbert space $\mathcal{H}$. It has an inner product which makes it possible to normalize all such vectors. Physical observables correspond to operators acting in Hilbert space. Chosing a basis, they can be represented by Hermetian matrices. The measured value of such an observable will be one of the eigenvalues of the corresponding matrix. This is one of the three basic postulates of quantum mechanics. As time evolves, the state vector for a system changes in a deterministic way according to the Schrödinger equation. Uncertainty in quantum mechanics comes in via the components of the state vector which are postulated to give the probability for measuring the corresponding eigenvalue.

### 2.1 Bra and ket vectors

It was shown by Dirac that a very elegant and compact formulation of the mathematics describing quantum systems is obtained when we introduce a new notation for the vectors and inner product in the previous chapter. A complex vector corresponding to z in Eq. (2.5) will now be written as $|\Psi\rangle \in \mathcal{H}$. Here $\Psi$ is just the name we have chosen for this particular vector, we could just as well kept the old name $z$ if we preferred that. Dirac called it a ket vector. The sum of two such vectors $\left|\Psi_{i}\right\rangle$ and $\left|\Psi_{j}\right\rangle$ with complex coefficients will be a new ket vector, i.e.

$$
\begin{equation*}
a\left|\Psi_{i}\right\rangle+b\left|\Psi_{j}\right\rangle=\left|\Psi_{k}\right\rangle \in \mathcal{H} \tag{2.1}
\end{equation*}
$$

There is also a dual space $\mathcal{H}^{*}$ where the vectors are called bra vectors and denoted by $\langle\Psi| \in \mathcal{H}^{*}$. We say that $\langle\Psi|$ is the adjoint of $|\Psi\rangle$ and corresponds to complex conjugation in the previous note where the dual vector was instead denoted by an upper star $*$. This operation is now indicated instead by an upper dagger $\dagger$. The adjoint of the linear combination (2.1) will therefore be

$$
\begin{equation*}
\left[a\left|\Psi_{i}\right\rangle+b\left|\Psi_{j}\right\rangle\right]^{\dagger}=a^{*}\left\langle\Psi_{i}\right|+b^{*}\left\langle\Psi_{j}\right| \in \mathcal{H}^{*} \tag{2.2}
\end{equation*}
$$

To convince yourself that this is right, rewrite it in the notation of Chapter 1 and it follows automatically.

For the inner product between two vectors $\left|\Psi_{i}\right\rangle$ and $\left|\Psi_{j}\right\rangle$ which we could write as $\left\langle\Psi_{i}\right| \cdot\left|\Psi_{j}\right\rangle$ in the old notation, Dirac wrote instead $\left\langle\Psi_{i} \mid \Psi_{j}\right\rangle$. We then also see why he introduced the names bra and ket vectors since when combined in the inner product, they form a bracket. The product has the following, basic properties

- $\left\langle\Psi_{i} \mid \Psi_{j}\right\rangle=\left\langle\Psi_{j} \mid \Psi_{i}\right\rangle^{*}$
- $\left\langle\Psi_{i} \mid \Psi_{i}\right\rangle \geq 0$
- $\left\langle\Psi \mid a \Psi_{i}+b \Psi_{j}\right\rangle=a\left\langle\Psi \mid \Psi_{i}\right\rangle+b\left\langle\Psi \mid \Psi_{j}\right\rangle$
which also follow directly from the definition of the inner product in the previous note. The squared norm or length of the vector $|\Psi\rangle$ is $\langle\Psi \mid \Psi\rangle$ and when this equals one, we say that the vector is normalized to one. When the inner product of the vectors $\left|\Psi_{i}\right\rangle$ and $\left|\Psi_{j}\right\rangle$ is zero, we say that they are orthogonal to each other.

Introducing a basis $|1\rangle,|2\rangle, \ldots,|N\rangle$ in the $N$-dimensional Hilbert space $\mathcal{H}_{N}$, the vectors $|\Psi\rangle$ can be decomposed in the standard way as

$$
\begin{equation*}
|\Psi\rangle=\sum_{n=1}^{N} \psi_{n}|n\rangle \tag{2.3}
\end{equation*}
$$

where the complex numbers $\psi_{n}$ are the components of the vector. When the basis vectors satisfy $\langle m \mid n\rangle=\delta_{m n}$, they form an orthonormal basis. The squared length of the vector is then

$$
\begin{equation*}
\langle\Psi \mid \Psi\rangle=\sum_{m, n=1}^{N} \psi_{m}^{*} \psi_{n}\langle m \mid n\rangle=\sum_{n=1}^{N} \psi_{n}^{*} \psi_{n} \tag{2.4}
\end{equation*}
$$

and is obviously positive.
Furthermore, from (2.3) we see that the components of the vector are given by the inner products

$$
\begin{equation*}
\psi_{n}=\langle n \mid \Psi\rangle \tag{2.5}
\end{equation*}
$$

Inserted back into (2.3), we then write it as

$$
|\Psi\rangle=\sum_{n=1}^{N} \psi_{n}|n\rangle=\sum_{n=1}^{N}|n\rangle\langle n \mid \Psi\rangle
$$

with the components following the basis vectors instead. Since this holds for all vectors $|\Psi\rangle$, the object multiplying it on the right must be a unit operator $\widehat{I}$ resulting from the completeness sum

$$
\begin{equation*}
\sum_{n=1}^{N}|n\rangle\langle n|=\widehat{I} \tag{2.6}
\end{equation*}
$$

such that for any vector $|\Psi\rangle$, we have $\widehat{I}|\Psi\rangle=|\Psi\rangle$. The completeness relation (2.6) has exactly the same structure as in the two previous notes, but here the order of the two vectors in the products of the sum is even more important than before. If they are interchanged, the products look like inner products which they are not.

### 2.2 Operators in Hilbert space

When the unit operator $\widehat{I}$ acts on a ket vector $|\Psi\rangle$, the result is the same vector. A general operator $\widehat{A}$ acting on $|\Psi\rangle$ is defined to give another vector $\left|\Psi^{\prime}\right\rangle$ in the same Hilbert space,

$$
\begin{equation*}
\widehat{A}|\Psi\rangle=\left|\Psi^{\prime}\right\rangle \in \mathcal{H} \tag{2.7}
\end{equation*}
$$

Operators can be linearly combined. For instance, if $\widehat{B}$ is another operator, then

$$
[a \widehat{A}+b \widehat{B}]|\Psi\rangle=a \widehat{A}|\Psi\rangle+b \widehat{B}|\Psi\rangle
$$

will also be a vector. There is also a zero operator $\widehat{0}$ which always gives the null vector as result, $\widehat{0}|\Psi\rangle=|0\rangle$.
The unit operator (2.6) is a linear combination of projection operators defined as

$$
\begin{equation*}
\widehat{\Pi}_{n}=|n\rangle\langle n| \tag{2.8}
\end{equation*}
$$

If such an operator acts on the general vector $|\Psi\rangle=\sum_{m} \psi_{m}|m\rangle$, we see that it gives a new vector which is the projection of $|\Psi\rangle$ along the basis vector $|n\rangle$,

$$
\begin{equation*}
\hat{\Pi}_{n}|\Psi\rangle=\psi_{n}|n\rangle \tag{2.9}
\end{equation*}
$$

and hence its name. When it acts once more on $|\Psi\rangle$, it just gives the same vector. We therefore have $\widehat{\Pi}_{n}^{2}=\widehat{\Pi}_{n}$ since $\widehat{\Pi}_{n}^{2}=|n\rangle\langle n \mid n\rangle\langle n|=|n\rangle\langle n|$. This can also be taken as the definition of a projection operator.

Not only projection operators, but all operators can be multiplied together. If the action of $\widehat{A}$ on $|\Psi\rangle$ in (2.7) is followed by the action of operator $\widehat{B}$, the result will be of the form

$$
\widehat{B} \widehat{A}|\Psi\rangle=\widehat{B}\left|\Psi^{\prime}\right\rangle=\left|\Psi^{\prime \prime}\right\rangle \in \mathcal{H}
$$

If the action of the operators $\widehat{A}$ and $\widehat{B}$ occured in the opposite order, then we would in general find another result, i.e. $\widehat{A} \widehat{B}|\Psi\rangle$ is usually different from the vector $\widehat{B} \widehat{A}|\Psi\rangle$. This is expressed by the special product combination called the commutator of $\widehat{A}$ and $\widehat{B}$

$$
\begin{equation*}
[\widehat{A}, \widehat{B}]=\widehat{A} \widehat{B}-\widehat{B} \widehat{A} \tag{2.10}
\end{equation*}
$$

which is then a non-zero operator.
For each operator $\widehat{A}$ there is an inverse operator denoted by $\widehat{A}^{-1}$ which is defined to give the unit operator when combined,

$$
\begin{equation*}
\widehat{A} \widehat{A}^{-1}=\widehat{A}^{-1} \widehat{A}=\widehat{I} \tag{2.11}
\end{equation*}
$$

The inverse $(\widehat{A} \widehat{B})^{-1}$ of a product $\widehat{A} \widehat{B}$ is the reversed product of inverses,

$$
\begin{equation*}
(\widehat{A} \widehat{B})^{-1}=\widehat{B}^{-1} \widehat{A}^{-1} \tag{2.12}
\end{equation*}
$$

since then

$$
(\widehat{A} \widehat{B})^{-1} \widehat{A} \widehat{B}=\widehat{B}^{-1} \widehat{A}^{-1} \widehat{A} \widehat{B}=\widehat{B}^{-1} \widehat{B}=\widehat{I}
$$

For the unit operator the inverse is obviously the unit operator itself.

### 2.3 Matrix representations

Given an orthonormal and complete $N$-dimensional basis $\{|n\rangle\}$, a general operator $\widehat{A}$ can be written as a linear combination of the elementary operators $|m\rangle\langle n|$,

$$
\begin{equation*}
\widehat{A}=\sum_{m, n}^{N} A_{m n}|m\rangle\langle n| \tag{2.13}
\end{equation*}
$$

The coefficients $A_{m n}$ are in general complex numbers and are seen to be given as

$$
\begin{equation*}
A_{m n}=\langle m| \widehat{A}|n\rangle \tag{2.14}
\end{equation*}
$$

We say that these numbers represent the abstract operator in this basis. They can be grouped into the $N \times N$ matrix

$$
A=\left(\begin{array}{cccc}
A_{11} & A_{12} & \ldots & A_{1 N}  \tag{2.15}\\
A_{21} & A_{22} & \ldots & A_{2 N} \\
\vdots & \vdots & \ddots & \vdots \\
A_{N 1} & A_{N 2} & \ldots & A_{N N}
\end{array}\right)
$$

which is then said to give a matrix representation of the operator.
The actual calculation of $\left|\Psi^{\prime}\right\rangle=\widehat{A}|\Psi\rangle$ is now most directly done by matrix calculation. We represent the abstract vector $|\Psi\rangle$ by its components $\psi_{n}=\langle n \mid \Psi\rangle$ which we group into the 1 -column matrix

$$
\psi=\left(\begin{array}{c}
\psi_{1} \\
\psi_{2} \\
\vdots \\
\psi_{N}
\end{array}\right)
$$

and similarly for $\left|\Psi^{\prime}\right\rangle$. It has the components

$$
\psi_{m}^{\prime}=\sum_{n=1}^{N} A_{m n} \psi_{n}
$$

which corresponds to the matrix equation $\psi^{\prime}=A \psi$ or

$$
\left(\begin{array}{c}
\psi_{1}^{\prime}  \tag{2.16}\\
\psi_{2}^{\prime} \\
\vdots \\
\psi_{N}^{\prime}
\end{array}\right)=\left(\begin{array}{cccc}
A_{11} & A_{12} & \ldots & A_{1 N} \\
A_{21} & A_{22} & \ldots & A_{2 N} \\
\vdots & \vdots & \ddots & \vdots \\
A_{N 1} & A_{N 2} & \ldots & A_{N N}
\end{array}\right)\left(\begin{array}{c}
\psi_{1} \\
\psi_{2} \\
\vdots \\
\psi_{N}
\end{array}\right)
$$

As a simple illustration, consider a 3 -dimensional Hilbert space with basis vectors $|1\rangle$, $|2\rangle$ and $|3\rangle$ in which there are the operators

$$
\widehat{A}=i|1\rangle\langle 3|+|2\rangle\langle 2|-i|3\rangle\langle 1|
$$

and

$$
\widehat{B}=2|1\rangle\langle 2|+2|2\rangle\langle 1|-3|3\rangle\langle 3|
$$

When acting upon the state vector

$$
|\Psi\rangle=3|1\rangle-2|2\rangle+|3\rangle
$$

we find $\widehat{A}|\Psi\rangle=i|1\rangle-2|2\rangle-3 i|3\rangle$ and $\widehat{B}|\Psi\rangle=-4|1\rangle+6|2\rangle-3|3\rangle$. To check if these two operators commute, we calculate $\widehat{A} \widehat{B}|\Psi\rangle=-3 i|1\rangle+6|2\rangle+4 i|3\rangle$. Since this is different from $\widehat{B} \widehat{A}|\Psi\rangle=-4|1\rangle+2 i|2\rangle+9 i|3\rangle$, we conclude that they have a non-vanishing commutator. It can in fact be calculated directly and we obtain

$$
[\widehat{A}, \widehat{B}]=2(|2\rangle\langle 1|-|1\rangle\langle 2|)-3 i(|1\rangle\langle 3|+|3\rangle\langle 1|)-2 i(|2\rangle\langle 3|+|3\rangle\langle 2|)
$$

which is certainly not the zero operator.
From the above discussion we can now also find the matrix representations of these two operators from using the formula (2.14) for the matrix elements. Or we can read them directly off the abstract forms of the operators which is faster. Either way, we find for the corresponding matrices

$$
A=\left(\begin{array}{rrr}
0 & 0 & i \\
0 & 1 & 0 \\
-i & 0 & 0
\end{array}\right), \quad B=\left(\begin{array}{rrr}
0 & 2 & 0 \\
2 & 0 & 0 \\
0 & 0 & -3
\end{array}\right)
$$

Since the operators $\widehat{A}$ and $\widehat{B}$ do not commute, we find equivalently that the matrices $A$ and $B$ do not commute under matrix multiplication.
In terms of matrices, the equivalent calculation of $\widehat{A}|\Psi\rangle=\left|\Psi^{\prime}\right\rangle$ will now be

$$
\left(\begin{array}{rrr}
0 & 0 & i \\
0 & 1 & 0 \\
-i & 0 & 0
\end{array}\right)\left(\begin{array}{r}
3 \\
-2 \\
1
\end{array}\right)=\left(\begin{array}{r}
i \\
-2 \\
-3 i
\end{array}\right)
$$

which give the components of the transformed vector in agreement with what we found above.

### 2.4 Adjoint operations

Previously we have called the bra vector $\langle\Psi|$ the dual or adjoint of the ket vector $|\Psi\rangle$. We can now introduce also the adjoint operator. Let $\left|\Psi^{\prime}\right\rangle=a|\Psi\rangle$ where $a$ is a complex number. The adjoint bra-vector is then $\left\langle\Psi^{\prime}\right|=a^{*}\langle\Psi|=\langle\Psi| a^{*}$ since $a^{*}$ is just a number. If now instead we have the vector $\left|\Psi^{\prime}\right\rangle=\widehat{A}|\Psi\rangle$ where $\widehat{A}$ is some operator, it is then natural to define the adjoint operator $\widehat{A}^{\dagger}$ by taking the adjoint of $\left|\Psi^{\prime}\right\rangle$, i.e.

$$
\begin{equation*}
\left\langle\Psi^{\prime}\right|=\langle\Psi| \widehat{A}^{\dagger} \tag{2.17}
\end{equation*}
$$

where $\widehat{A}^{\dagger}$ here acts to the left. Thus we also have that $(a \widehat{A})^{\dagger}=a^{*} \widehat{A}^{\dagger}$.
Matrix elements of the adjoint operator are given by the inner product $\left\langle\Psi_{j}\right| \widehat{A}^{\dagger}\left|\Psi_{i}\right\rangle$. If we here introduce $\left|\Psi_{k}\right\rangle=\widehat{A}\left|\Psi_{j}\right\rangle$, we then have the important result

$$
\begin{equation*}
\left\langle\Psi_{j}\right| \hat{A}^{\dagger}\left|\Psi_{i}\right\rangle=\left\langle\Psi_{k} \mid \Psi_{i}\right\rangle=\left\langle\Psi_{i} \mid \Psi_{k}\right\rangle^{*}=\left\langle\Psi_{i}\right| \widehat{A}\left|\Psi_{j}\right\rangle^{*} \tag{2.18}
\end{equation*}
$$

In an ortonormal basis we then have the matrix elements $A_{m n}^{\dagger}=\langle m| \widehat{A}^{\dagger}|n\rangle=\langle n| \widehat{A}|m\rangle^{*}=$ $A_{n m}^{*}=A_{m n}^{T *}$. The adjoint operator is therefore represented by the adjoint or Hermitean conjagated matrix $A^{\dagger}=A^{T *}$ first defined by (1.33) in the previous chapter. Letting $\widehat{A} \rightarrow \widehat{A}^{\dagger}$ in (2.18) it follows also that $\widehat{A}^{\dagger \dagger}=\widehat{A}$.

The adjoint of a product of operators is the reversed product of the adjoint operators, e.g. for two operators,

$$
\begin{equation*}
(\widehat{A} \widehat{B})^{\dagger}=\widehat{B}^{\dagger} \widehat{A}^{\dagger} \tag{2.19}
\end{equation*}
$$

This corresponds to (2.12) for the inverse of the product. It follows again from the definition (2.17) of the adjoint operator by considering

$$
\left\langle\Psi_{i}\right| \widehat{A} \widehat{B}\left|\Psi_{j}\right\rangle^{*}=\left\langle\Psi_{j}\right|(\widehat{A} \widehat{B})^{\dagger}\left|\Psi_{i}\right\rangle
$$

Now write $\left|\Psi_{k}\right\rangle=\widehat{B}\left|\Psi_{j}\right\rangle$, use $\left\langle\Psi_{i}\right| \widehat{A}\left|\Psi_{k}\right\rangle^{*}=\left\langle\Psi_{k}\right| \widehat{A}^{\dagger}\left|\Psi_{i}\right\rangle$ which equals $\left\langle\Psi_{j}\right| \widehat{B}^{\dagger} \widehat{A}^{\dagger}\left|\Psi_{i}\right\rangle$ and we have the stated result.

A very important class of operators are those who have an adjoint which equals the operator itself, i.e.

$$
\begin{equation*}
\hat{A}^{\dagger}=\widehat{A} \tag{2.20}
\end{equation*}
$$

They are said to be self-adjoint or Hermitian. The matrix elements (2.14) of such an operator will therefore have the property that

$$
\begin{equation*}
A_{m n}=\langle m| \widehat{A}|n\rangle=\langle n| \widehat{A}|m\rangle^{*}=A_{n m}^{*}=A_{m n}^{\dagger} \tag{2.21}
\end{equation*}
$$

and is therefore said to be Hermitean or self-adjoint.
We see from this that diagonal matrix elements $A_{n n}=\langle n| \widehat{A}|n\rangle$ of a Hermitian operator are always real numbers, or more generally $\langle\Psi| \widehat{A}|\Psi\rangle=\langle\Psi| \widehat{A}|\Psi\rangle^{*}$ for any vector $|\Psi\rangle$.

### 2.5 Eigenvectors and eigenvalues

When an operator $\widehat{A}$ acts on some vector $|\Psi\rangle$, one usually obtains a new vector $\left|\Psi^{\prime}\right\rangle$. But one can easily imagine the special situation where the result is the same vector or a vector proportional to the initial vector $|\Psi\rangle$. Then we have

$$
\begin{equation*}
\widehat{A}|\Psi\rangle=a|\Psi\rangle \tag{2.22}
\end{equation*}
$$

where the number $a$ is general is complex. We then say that $|\Psi\rangle$ is an eigenvector for the operator $\widehat{A}$ and $a$ is the corresponding eigenvalue.

The eigenvalues for Hermitian operators are real. This follows from the eigenvalue equation (2.22) by multiplying with the bra vector $\langle\Psi|$ in from the left, giving

$$
\langle\Psi| \widehat{A}|\Psi\rangle=a\langle\Psi \mid \Psi\rangle
$$

Now we have just seen that the left-hand side $\langle\Psi| \widehat{A}|\Psi\rangle$ is a real number. Then the eigenvalue $a$ must also be real since the norm $\langle\Psi \mid \Psi\rangle$ is real.

An operator can in general have many different eigenvectors. The eigenvectors of an Hermitian operator $\widehat{A}$ belonging to different eiegnvalues are orthogonal to each other. To demonstarte this very important property, consider the two eigenvectors $\left|\Psi_{m}\right\rangle$ and $\left|\Psi_{n}\right\rangle$ with the corresponding eigenvalues $a_{m}$ and $a_{n}$,

$$
\widehat{A}\left|\Psi_{m}\right\rangle=a_{m}\left|\Psi_{m}\right\rangle, \quad \widehat{A}\left|\Psi_{n}\right\rangle=a_{n}\left|\Psi_{n}\right\rangle
$$

We multiply the first equation with $\left\langle\Psi_{n}\right|$ from the left and then use $\left\langle\Psi_{n}\right| \widehat{A}=a_{n}\left\langle\Psi_{n}\right|$ which follows from taking the adjoint of the second equation. It then follows that $a_{m}\left\langle\Psi_{n} \mid \Psi_{m}\right\rangle=a_{n}\left\langle\Psi_{n} \mid \Psi_{m}\right\rangle$ or

$$
\begin{equation*}
\left(a_{m}-a_{n}\right)\left\langle\Psi_{n} \mid \Psi_{m}\right\rangle=0 \tag{2.23}
\end{equation*}
$$

So, if $a_{m} \neq a_{n}$, we must have $\left\langle\Psi_{n} \mid \Psi_{m}\right\rangle=0$ and the eigenvectors are orthogonal.
If the eigenvalues $a_{m}$ and $a_{n}$ belonging to different eigenvectors happen to be equal, then we say that they are degenerate. Then it is no longer necessary that the eigenvectors satisfy $\left\langle\Psi_{n} \mid \Psi_{m}\right\rangle=0$ and they are in general not orthogonal to each other. But by taking linear combinations of these degenerate eigenvectors, we can always arrange it so that even in this case all the new eigenvectors are orthogonal to each other.

An important mathematical theorem, which we shall not try to prove, says that all the eigenvectors of an Hermitian operator can be used to form a complete set of orthonormal basis vectors in the corresponding Hilbert space. If the dimension of this space is $N$, then these orthonormalized eigenvectors will therefore satisfy the completeness relation (2.6). Are there other Hermitian operators defined in the same Hilbert space, then their eigenvectors will also form complete sets and can be used as basis vectors in the same Hilbert space.
When a vector $|\Psi\rangle$ is simultaneously an eigenvector of two operators $\widehat{A}$ and $\widehat{B}$, these operators must then commute with each other. This is easy to show. In addition to (2.22), we then have the condition $\widehat{B}|\Psi\rangle=b|\Psi\rangle$. Operating on this with $\widehat{A}$, we get $\widehat{A} \widehat{B}|\Psi\rangle=b \widehat{A}|\Psi\rangle=a b|\Psi\rangle$. Similarly, we find $\widehat{B} \widehat{A}|\Psi\rangle=a b|\Psi\rangle$. Subtracting these two results from each other, it then follows that $(\widehat{A} \widehat{B}-\widehat{B} \widehat{A})|\Psi\rangle=0$. This is now satisfied for all eigenstates $|\Psi\rangle$ when $[\widehat{A}, \widehat{B}]=0$. The reverse statement has great practical use, i.e. when two operators commute with each other, we can find eigenvectors which are simultaneously eigenvectors of both operators.

### 2.6 The quantum postulates

Quantum mechanics cannot be derived from classical mechanics or any other branch of physics. If that had been the case, quantum mechanics would be nothing new. It is based on new ideas and must be formulated in a new language.

In the mathematical language of bra and ket evctors, the basic ingredients of quantum mechanics can be stated in terms of the following three axioms:

1. For each physical observable $A$, there is a corresponding Hermitian operator $\widehat{A}$. Measured values of the observable will be the eigenvalues $a_{n}$ of the operator where

$$
\begin{equation*}
\widehat{A}|n\rangle=a_{n}|n\rangle \tag{2.24}
\end{equation*}
$$

for the corresponding eigenvector $|n\rangle$. The energy is given by the eigenvalues of the Hamiltonian operator $\widehat{H}$.
2. The observable properties of a physical system is at any time coded into a state vector $|\Psi\rangle=|\Psi, t\rangle$ in the Hilbert space for the system. The expectation value of the observable $A$ when the system is in this state, is given by

$$
\begin{equation*}
\langle\widehat{A}\rangle=\langle\Psi| \widehat{A}|\Psi\rangle \tag{2.25}
\end{equation*}
$$

when the state vector is normalized to have unit norm, i.e. $\langle\Psi \mid \Psi\rangle=1$.
3. The state vector $|\Psi, t\rangle$ varies with time $t$ according to the Schrödinger equation

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t}|\Psi, t\rangle=\widehat{H}|\Psi, t\rangle \tag{2.26}
\end{equation*}
$$

where $\widehat{H}$ is the Hamiltonian operator.
The only new parameter which is needed, is the Planck-Dirac constant $\hbar=h / 2 \pi$ where $h$ is the original Planck constant.

To see a little of what this implies, let us express the state vector $|\Psi\rangle$ by its components in the basis formed by the eigenvectors of $\widehat{A}$. Since these form a complete, orthonormal set of vectors, we have as before

$$
|\Psi\rangle=\sum_{n} \psi_{n}|n\rangle
$$

where the components $\psi_{n}=\langle n \mid \Psi\rangle$ and $\widehat{A}|n\rangle=a_{n}|n\rangle$. Since we assume that $|\Psi\rangle$ has unit norm, the components satify the condition

$$
\begin{equation*}
\sum_{n}\left|\psi_{n}\right|^{2}=1 \tag{2.27}
\end{equation*}
$$

The expectation value of the quantum observable $\widehat{A}$ is then

$$
\langle\widehat{A}\rangle=\sum_{n} \psi_{n}\langle\Psi| \widehat{A}|n\rangle=\sum_{n} \psi_{n} a_{n} \psi_{n}^{*}=\sum_{n} a_{n}\left|\psi_{n}\right|^{2}
$$

So the average value of the measurement is the weighted average of all the eigenvalues of $\widehat{A}$. The probability to measure eigenvalue $a_{n}$ is seen to be $P_{n}=\left|\psi_{n}\right|^{2}$. From the unit norm of $|\Psi\rangle$ these probabilities sum up to one as they should. It is therefore natural to call the complex number $\psi_{n}=\langle n \mid \Psi\rangle$ the probability amplitude to find this measured value. Immediately after the measurement giving the eigenvalue $a_{n}$, the system is with certainty in the state $|n\rangle$. We say that the state vector has collapsed from $|\Psi\rangle$ to $|n\rangle$.
This probability interpretation means that a measurement of the average value $\langle\hat{A}\rangle$, also called the expectation value of the operator $\widehat{A}$, really involves an infinite number of single measurements. The first measurement gives a certain eigenvalue. Then the measurement is repeated on an identical system prepared in the same state $|\Psi\rangle$. This gives in general a different eigennvalue as result. Continuing this way on a large number of identical systems in the same state, we then calculate the average value of all these single measurements and thus get the expectation value $\langle\widehat{A}\rangle$.

### 2.7 Ehrenfest's theorem

Just as a classical observable $A(t)$ can explicitly depend on time, the corresponding quantum operator $\widehat{A}(t)$ can also be time dependent. The expectation value

$$
\begin{equation*}
\langle\widehat{A}, t\rangle=\langle\Psi, t| \widehat{A}(t)|\Psi, t\rangle \tag{2.28}
\end{equation*}
$$

therefore gets a time dependence both from $\widehat{A}(t)$, from $|\Psi, t\rangle$ and from $\langle\Psi, t|$. From the Schrödinger equation (2.26) we know how the ket vector $|\Psi, t\rangle$ varies with time. Taking the adjoint,

$$
\begin{equation*}
-i \hbar \frac{\partial}{\partial t}\langle\Psi, t|=\langle\Psi, t| \widehat{H} \tag{2.29}
\end{equation*}
$$

we also can find how the corresponding bra vector varies. This now enables us to calculate the time derivative of the full expectation value $\langle A(t)\rangle$ which is

$$
\frac{d}{d t}\langle\widehat{A}, t\rangle=\left(\frac{\partial}{\partial t}\langle\Psi, t|\right) \widehat{A}(t)|\Psi, t\rangle+\langle\Psi, t| \widehat{A}(t)\left(\frac{\partial}{\partial t}|\Psi, t\rangle\right)+\langle\Psi, t|\left(\frac{\partial}{\partial t} \widehat{A}(t)\right)|\Psi, t\rangle
$$

Now using the Schrödinger equation and its adjoint, the two first terms can be expressed by the action of the Hamiltonian operator with the result

$$
\begin{equation*}
\frac{d}{d t}\langle\widehat{A}, t\rangle=\frac{i}{\hbar}\langle\Psi|[\widehat{H}, \widehat{A}]|\Psi\rangle+\langle\Psi| \frac{\partial \widehat{A}}{\partial t}|\Psi\rangle \tag{2.30}
\end{equation*}
$$

This is called Ehrenfest's theorem since the result on the right for the expectation value bears a very close resemblance to what would be the result in classical mechanics. This will become more clearer later. At this stage we can just say that the physical content of this theorem, which we can write in a bit more compact form as

$$
\begin{equation*}
\frac{d}{d t}\langle\widehat{A}, t\rangle=\frac{i}{\hbar}\langle[\widehat{H}, \widehat{A}]\rangle+\left\langle\frac{\partial \widehat{A}}{\partial t}\right\rangle, \tag{2.31}
\end{equation*}
$$

is to show that the time variation of a quantum mechanical expectation value is the same as the time variation of the corresponding classical variable.
When the operator $\widehat{A}$ has no explicit time dependence and also commutes with the Hamiltonian operator, we see that the expectation value $\langle\widehat{A}\rangle$ is constant in time. It is therefore said to be conserved. We will later see that such an operator corresponds to a symmetry of the physical system.

### 2.8 The time evolution operator

The Hamiltonian operator $\widehat{H}$ can in general depend on time. When it does not, we say we have a stationary system. From Ehrenfest's theorem we then see that the average energy of the system is constant, or to say it equivalently, it is conserved in time. In this important case the Schrödinger equation (2.26) can then easily be integrated to give

$$
\begin{equation*}
|\Psi, t\rangle=e^{-i \widehat{H} t / \hbar}|\Psi\rangle \tag{2.32}
\end{equation*}
$$

where $|\Psi\rangle=|\Psi, t=0\rangle$. The exponential of the Hamiltonian is meant to mean the standard expansion

$$
\begin{equation*}
e^{-i \widehat{H} t / \hbar}=\widehat{I}+\frac{1}{1!}\left(-\frac{i}{\hbar} \widehat{H} t\right)+\frac{1}{2!}\left(-\frac{i}{\hbar} \widehat{H} t\right)^{2}+\ldots \tag{2.33}
\end{equation*}
$$

Differentiating term by term, we verify that the derivative of the exponential is the expected result

$$
\begin{equation*}
\frac{\partial}{\partial t} e^{-i \widehat{H} t / \hbar}=-\frac{i}{\hbar} \widehat{H} e^{-i \widehat{H} t / \hbar} \tag{2.34}
\end{equation*}
$$

This also shows that (2.32) is a formal solution of the Schrödinger equation,

$$
i \hbar \frac{\partial}{\partial t}|\Psi, t\rangle=\widehat{H} e^{-i \widehat{H} t / \hbar}|\Psi\rangle=\widehat{H}|\Psi, t\rangle
$$

which has great practical use.
From the above we see that that the time evolution of the quantum system lies in the time development operator

$$
\begin{equation*}
\widehat{U}(t)=e^{-i \widehat{H} t / \hbar} \tag{2.35}
\end{equation*}
$$

which moves the system forward in time. The inverse operator

$$
\begin{equation*}
\widehat{U}^{-1}(t)=e^{i \widehat{H} t / \hbar}=\widehat{U}(-t) \tag{2.36}
\end{equation*}
$$

moves the system correspondingly back in time. Since the Hamiltonian operator is Hermitian, we see that the adjoint of the time development operator equals it inverse,

$$
\begin{equation*}
\widehat{U}^{\dagger}(t)=e^{i \widehat{H} t / \hbar}=\widehat{U}^{-1}(t) \tag{2.37}
\end{equation*}
$$

or $\hat{U} \hat{U}^{\dagger}=\widehat{U}^{\dagger} \widehat{U}=\hat{I}$. Such operators are said to be unitary. In fact, the time development operator forms the unitary Lie group $U(1)$ discussed in Note 2 . Each element $\widehat{U}(t)$ has an inverse $\widehat{U}^{-1}(t)$, the unit element is $\widehat{U}(0)=\widehat{I}$ and the product of two time developments is a third, $\widehat{U}\left(t_{1}\right) \widehat{U}\left(t_{2}\right)=\widehat{U}\left(t_{3}\right)$ where $t_{3}=t_{1}+t_{2}$.
A short time $\Delta t$ after the initial time $t=0$, the state vector has changed into

$$
|\Psi, \Delta t\rangle=\hat{U}(\Delta t)|\Psi\rangle
$$

where

$$
\widehat{U}(\Delta t)=\widehat{I}-\frac{i}{\hbar} \widehat{H} \Delta t+\mathcal{O}(\Delta t)^{2}
$$

and is caused by the Hamiltonian operator. We say that it is the generator of time development. From the group property of the operator, we can now build up an evolution over finite time $t=N \Delta t$ as a product of $N$ short ones, each of duration $\Delta t=t / N \rightarrow 0$. Thus we can write $\widehat{U}(t)=\widehat{U}(\Delta t) \widehat{U}(\Delta t) \cdots \widehat{U}(\Delta t)$ or

$$
\widehat{U}(t)=\left(\widehat{I}-\frac{i}{\hbar} \widehat{H} \frac{t}{N}\right)^{N}
$$

in the limit $N \rightarrow \infty$. Now making use of the definition of the exponential function,

$$
\begin{equation*}
e^{x}=\left(1+\frac{x}{N}\right)^{N} \tag{2.38}
\end{equation*}
$$

when $N \rightarrow \infty$, we see that we are back at the finite expression (2.35) for the time development operator.

### 2.9 Stationary states and energy basis

A stationary state $|E, t\rangle$ is defined to have a simple exponential time dependence,

$$
|E, t\rangle=|E\rangle e^{-i E t / \hbar}
$$

where $|E\rangle$ is the state vector at time $t=0$. Using this in the time-dependent Schrödinger equation (2.26), we see that is must satisfy

$$
\begin{equation*}
\widehat{H}|E\rangle=E|E\rangle \tag{2.39}
\end{equation*}
$$

The stationary states are therefore the eigenvectors of the Hamiltonian with $E$ being the corresponding energy eigenvalue.
In a given ortonormal basis $\{|n\rangle\}$ the eigenvectors in (2.39) will be of the form

$$
|E\rangle=\sum_{n=1}^{\infty} u_{n}|n\rangle
$$

with the amplitudes $u_{n}=\langle n \mid E\rangle$. As in the previous note we can then write the eigenvalue eqaution on matrix form as $H u=E u$ where $H$ is the Hamiltonian matrix in this basis and $u$ is the 1 -column matrix representing the vector $|E\rangle$. More explicitly, it becomes

$$
\left(\begin{array}{cccc}
H_{11} & H_{12} & \ldots & H_{1 N}  \tag{2.40}\\
H_{21} & H_{22} & \ldots & H_{2 N} \\
\vdots & \vdots & \ddots & \vdots \\
H_{N 1} & H_{N 2} & \ldots & H_{N N}
\end{array}\right)\left(\begin{array}{c}
u_{1} \\
u_{2} \\
\vdots \\
u_{N}
\end{array}\right)=E\left(\begin{array}{c}
u_{1} \\
u_{2} \\
\vdots \\
u_{N}
\end{array}\right)
$$

We thus have $N$ equations for the unknown amplitudes $u_{n}$. Since this set of equations is homogenous, we will only have non-trivial solutions when the determinant formed by their coefficient is zero, i.e.

$$
\left|\begin{array}{cccc}
H_{11}-E & H_{12} & \ldots & H_{1 N}  \tag{2.41}\\
H_{21} & H_{22}-E & \ldots & H_{2 N} \\
\vdots & \vdots & \ddots & \vdots \\
H_{N 1} & H_{N 2} & \ldots & H_{N N}-E
\end{array}\right|=0
$$

Writing this out, we find a polynomial equation of $N$ th order. The energy eigenvalues are then the zeros of the equation. Since the Hamiltonian is Hermitian, we are guaranteed that they are all on the real axis. For each such eigenvalue $E_{n}$ we can then solve for the $N$ amplitudes $u_{n}$ which then forms the corresponding eigenvector $\left|E_{n}\right\rangle$.

It is most convenient to normalize all the eigenvectors to have length equal to one. Different eigenvalues will have ortogonal eigenvectors. If some eigenvalues turn out to be equal, they are said to be degenerate. The corresponding eigenvectors are not necessarily automatically orthogonal to each other. But by taking linear combinations, this can be achieved. In the end we then have a complete set

$$
\sum_{n=1}^{N}\left|E_{n}\right\rangle\left\langle E_{n}\right|=\widehat{I}
$$

of ortonormal eigenvectors $\left\langle E_{m} \mid E_{n}\right\rangle=\delta_{m n}$ which provides us with a new basis in the Hilbert space we are working.

This new basis is particularly useful for describing the time evolution of an arbitrary state vector $|\Psi\rangle$. Since we now can write

$$
|\Psi\rangle=\sum_{n=1}^{N} C_{n}\left|E_{n}\right\rangle
$$

with $C_{n}=\left\langle E_{n} \mid \Psi\right\rangle$, we can use the time development operator (2.35) to find the corresponding new state $|\Psi, t\rangle$ at a later time. It gives

$$
\begin{align*}
|\Psi, t\rangle & =e^{-i \widehat{H} t / \hbar}|\Psi\rangle=\sum_{n=1}^{N} C_{n} e^{-i \widehat{H} t / \hbar}\left|E_{n}\right\rangle \\
& =\sum_{n=1}^{N} C_{n} e^{-i E_{n} t / \hbar}\left|E_{n}\right\rangle \equiv \sum_{n=1}^{N} C_{n}(t)\left|E_{n}\right\rangle \tag{2.42}
\end{align*}
$$

which can now be used to answer specific questions about the properties of the system as time goes by. For instance, the probability amplitude that this system is found in one of the states $|m\rangle$ of the original basis set is

$$
\begin{equation*}
\psi_{m}(t)=\langle m \mid \Psi, t\rangle=\sum_{n=1}^{N} C_{n} e^{-i E_{n} t / \hbar}\left\langle m \mid E_{n}\right\rangle \tag{2.43}
\end{equation*}
$$

depending on the components $\left\langle m \mid E_{n}\right\rangle$ of the energy basis vectors onto the original basis vectors.

### 2.10 Change of basis

In general we can have several basis sets to use for matrix calculations in Hilbert space. Typically we have a a position basis $\{|n\rangle\}$ corresponding to the available positions of a particle in the system and also an energy basis provided by the eigenvectors $\left\{\left|E_{n}\right\rangle\right\}$ of the Hamiltonian $\widehat{H}$. Operators and state vectors will have different matrix representations in these two bases. For example, in the position basis the Hamiltonian is generally given by a non-diagonal matrix while in the energy basis it is diagonal with matrix elements $H_{m n}=\left\langle E_{m}\right| \widehat{H}\left|E_{n}\right\rangle=E_{m} \delta_{m n}$.
Let us assume more generally that we have two orthonormal and complete basis sets $\{|n\rangle\}$ and $\left\{\left|n^{\prime}\right\rangle\right\}$. A state vector $|\psi\rangle$ is then expressed as

$$
|\psi\rangle=\sum_{n} \psi_{n}|n\rangle
$$

with components $\psi_{n}=\langle n \mid \psi\rangle$ or

$$
|\psi\rangle=\sum_{n^{\prime}} \psi_{n^{\prime}}\left|n^{\prime}\right\rangle
$$

Here the components are $\psi_{n^{\prime}}=\left\langle n^{\prime} \mid \psi\right\rangle$ which can be written as

$$
\psi_{n^{\prime}}=\sum_{n}\left\langle n^{\prime} \mid n\right\rangle\langle n \mid \psi\rangle=\sum_{n} U_{n^{\prime} n} \psi_{n}
$$

More compactly we write the result as the matrix product

$$
\begin{equation*}
\psi^{\prime}=U \psi \tag{2.44}
\end{equation*}
$$

where $\psi^{\prime}$ is the one-column matrix with components $\psi_{n^{\prime}}$ representing the vector $\left|\psi^{\prime}\right\rangle$ and similarly for $\psi$. This is the wanted transformation of the vector components from one basis to the other. It is given by the transformation matrix $U=\left(U_{n^{\prime} n}\right)$ with matrix elements $U_{n^{\prime} n}=\left\langle n^{\prime} \mid n\right\rangle$.

Since both basis sets are orthonormal, the transformation matrix $U$ is unitary. This follows from

$$
\langle m \mid n\rangle=\delta_{m n}=\sum_{n^{\prime}}\left\langle m \mid n^{\prime}\right\rangle\left\langle n^{\prime} \mid n\right\rangle=\sum_{n^{\prime}} U_{n^{\prime} m}^{*} U_{n^{\prime} n}=\sum_{n^{\prime}} U_{m n^{\prime}}^{\dagger} U_{n^{\prime} n}
$$

or $U^{\dagger} U=I$ where the adjoint matrix is $U^{\dagger}=U^{* T}$. Similarly, from $\left\langle m^{\prime} \mid n^{\prime}\right\rangle=\delta_{m^{\prime} n^{\prime}}$ follows $U U^{\dagger}=I$. The inverse matrix $U^{-1}=\left(U_{n n^{\prime}}^{-1}\right)$ thus has matrix elements $U_{n n^{\prime}}^{-1}=$ $\left\langle n \mid n^{\prime}\right\rangle$ so that $U^{-1}=U^{\dagger}$. All this is just a repetion of linear transformations in complex vector spaces as already discussed in Chapter 1.
An operator $\widehat{A}$ is represented by the matrix $A=\left(A_{m n}\right)$ with elements $A_{m n}=\langle m| \widehat{A}|n\rangle$. Correspondingly, in the primed basis is is given by the matrix $A^{\prime}$ with elements $A_{m^{\prime} n^{\prime}}=$ $\left\langle m^{\prime}\right| \widehat{A}\left|n^{\prime}\right\rangle$. Inserting here un-primed basis sets, we find

$$
A_{m^{\prime} n^{\prime}}=\sum_{m n}\left\langle m^{\prime} \mid m\right\rangle A_{m n}\left\langle n \mid n^{\prime}\right\rangle=\sum_{m n} U_{m^{\prime} m} A_{m n} U_{n n^{\prime}}^{-1}
$$

which in matrix notation is

$$
\begin{equation*}
A^{\prime}=U A U^{-1} \tag{2.45}
\end{equation*}
$$

In mathematics this is called a similarity transformation of the matrix $A$.

### 2.11 Schrödinger and Heisenberg pictures

So far we have pictured the basis vectors as fixed in Hilbert space while state vectors will in general move with time. This is called the Schrödinger picture. But since it is only the relative motion betwen basis vectors and state vectors that matters, we are also allowed to consider the so-called Heisenberg picture where the basis vectors move so that the state vectors stay fixed. When the Hamiltonian operator $\widehat{H}$ has no explicit time dependence, we already know that the state vector varies with time as

$$
\begin{equation*}
|\Psi, t\rangle=\widehat{U}(t)|\Psi\rangle \tag{2.46}
\end{equation*}
$$

in the Schrödinger picture. Here $\widehat{U}(t)=\exp (-i \widehat{H} t / \hbar)$ is the time development operator already introduced. The constant vector $|\Psi\rangle$ can then be taken to be the state vector in the Heisenberg picture, i.e. $|\Psi\rangle_{H}=|\Psi\rangle$. If we now consider an observable corresponding to the operator $\widehat{A}$, the expectation value $\langle\widehat{A}\rangle=\langle\Psi, t| \widehat{A}|\Psi, t\rangle$ is the value of physical measurements and must be independent of which picture we make use of. Since

$$
\langle\widehat{A}\rangle=\langle\Psi| e^{i \widehat{H} t / \hbar} \widehat{A} e^{-i \widehat{H} t / \hbar}|\Psi\rangle
$$

we can then write it as $\langle\widehat{A}\rangle={ }_{H}\langle\Psi| \widehat{A}_{H}(t)|\Psi\rangle_{H}$ where now

$$
\begin{equation*}
\widehat{A}_{H}(t)=e^{i \widehat{H} t / \hbar} \widehat{A} e^{-i \widehat{H} t / \hbar} \tag{2.47}
\end{equation*}
$$

is the operator in the Heisenberg picture. This is like the above transformations between different representations. But here it is the time development operator $\widehat{U}(t)$ which provides the transformations of states and operators from one picture to the other. The Hamiltonian operator itself is seen to be the same in the two pictures.

Taking the time derivative of the Heisenberg operator (2.47) we find

$$
\begin{align*}
\frac{d}{d t} \widehat{A}_{H}(t) & =\frac{i}{\hbar} e^{i \widehat{H} t / \hbar} \widehat{H} \widehat{A} e^{-i \widehat{H} t / \hbar}-\frac{i}{\hbar} e^{i \widehat{H} t / \hbar} \widehat{A} \widehat{H} e^{-i \widehat{H} t / \hbar} \\
& =\frac{i}{\hbar}\left[\widehat{H}, \widehat{A}_{H}\right] \tag{2.48}
\end{align*}
$$

assuming there is no explicit time dependence in the operator $\widehat{A}$. This is the equation of motion for operators in the Heisenberg picture. Later we will see that the resulting equations have exactly the same structure as in classical mechanics. In fact, taking the expectation value of the equation, we recover Ehrenfest's theorem as derived earlier.
In the above we have taken the derivative of an operator. Formally it is defined here in the same way as in ordinary calculus, i.e.

$$
\begin{equation*}
\frac{d}{d t} \widehat{A}(t)=\lim _{\varepsilon \rightarrow 0} \frac{1}{\varepsilon}[\widehat{A}(t+\varepsilon)-\widehat{A}(t)] \tag{2.49}
\end{equation*}
$$

For example, if $\widehat{A}(t)=e^{t \widehat{B}}$ then $\widehat{A}(t+\varepsilon)=e^{(t+\varepsilon) \widehat{B}}=e^{t \widehat{B}}(\widehat{I}+\varepsilon \widehat{B}+\ldots)$ and thus

$$
\frac{d}{d t} \widehat{A}(t)=\widehat{B} e^{t \widehat{B}}=e^{t \widehat{B}} \widehat{B}
$$

as expected and already used. It also follows from expanding the exponential and taking the derivative of each term as we did the first time we encountered an exponential of an operator. From the same prescription one also obtains results like

$$
\frac{d}{d t}(\widehat{A}(t) \widehat{B}(t))=\widehat{A} \frac{d \widehat{B}}{d t}+\frac{d \widehat{A}}{d t} \widehat{B}
$$

and

$$
\frac{d}{d t} \widehat{A}^{-1}(t)=-\widehat{A}^{-1} \frac{d \widehat{A}}{d t} \widehat{A}^{-1}
$$

Notice that an operator and its derivative don't in general commute with each other.
Calculating with operators can be difficult since they generally don't commute, Some simplifications can often be made using the formulas

$$
\begin{equation*}
[\widehat{X}+\widehat{Y}, \widehat{Z}]=[\widehat{X}, \widehat{Z}]+[\widehat{Y}, \widehat{Z}] \tag{2.50}
\end{equation*}
$$

and

$$
\begin{equation*}
[\widehat{X} \widehat{Y}, \widehat{Z}]=\widehat{X}[\widehat{Y}, \widehat{Z}]+[\widehat{X}, \widehat{Z}] \widehat{Y} \tag{2.51}
\end{equation*}
$$

Both are easily proven by explicitly writing out both sides of the equations using the definition of the commutator. For example, (2.50) gives

$$
\begin{aligned}
{[\widehat{X}+\widehat{Y}, \widehat{Z}] } & =(\widehat{X}+\widehat{Y}) \widehat{Z}-\widehat{Z}(\widehat{X}+\widehat{Y})=\widehat{X} \widehat{Z}+\widehat{Y} \widehat{Z}-\widehat{Z} \widehat{X}-\widehat{Z} \widehat{Y} \\
& =\widehat{X} \widehat{Z}-\widehat{Z} \widehat{X}+\widehat{Y} \widehat{Z}-\widehat{Z} \widehat{Y}=[\widehat{X}, \widehat{Z}]+[\widehat{Y}, \widehat{Z}]
\end{aligned}
$$

The important Jacobi identity

$$
\begin{equation*}
[\widehat{X},[\hat{Y}, \widehat{Z}]]+[\widehat{Y},[\widehat{Z}, \widehat{X}]]+[\widehat{Z},[\widehat{X}, \widehat{Y}]]=0 \tag{2.52}
\end{equation*}
$$

is also easily verified that way.

### 2.12 The Lie formula

More cumbersome is the evaluation of an expression on the form $e^{\widehat{X}} \widehat{Y} e^{-\widehat{X}}$ appearing in the previous note when transforming an operator to the Heisenberg picture. We will now show that it is given by the Lie formula

$$
\begin{equation*}
e^{\widehat{X}} \widehat{Y} e^{-\widehat{X}}=\widehat{Y}+[\widehat{X}, \widehat{Y}]+\frac{1}{2!}[\widehat{X},[\widehat{X}, \widehat{Y}]]+\frac{1}{3!}[\widehat{X},[\widehat{X},[\widehat{X}, \widehat{Y}]]]+\ldots \tag{2.53}
\end{equation*}
$$

To prove it, consider the operator function

$$
\widehat{F}(t)=e^{t \widehat{X}} \widehat{Y} e^{-t \widehat{X}}
$$

It has the first derivative

$$
\frac{d \widehat{F}}{d t}=e^{t \widehat{X}} \widehat{X} \widehat{Y} e^{-t \widehat{X}}-e^{t \widehat{X}} \widehat{Y} \widehat{X} e^{-t \widehat{X}}=\widehat{X} \widehat{F}-\widehat{F} \widehat{X}=[\widehat{X}, \widehat{F}]
$$

The second derivative is similarly

$$
\frac{d^{2} \widehat{F}}{d t^{2}}=\widehat{X} \frac{d \widehat{F}}{d t}-\frac{d \widehat{F}}{d t} \widehat{X}=\left[\widehat{X}, \frac{d \widehat{F}}{d t}\right]=[\widehat{X},[\widehat{X}, \widehat{F}]]
$$

and so on. From the Taylor-expansion

$$
\widehat{F}(t)=\widehat{F}(0)+\frac{t}{1!} \widehat{F}^{\prime}(0)+\frac{t^{2}}{2!} \widehat{F}^{\prime \prime}(0)+\ldots
$$

we now get

$$
\widehat{F}(t)=\widehat{Y}+\frac{t}{1!}[\widehat{X}, \widehat{Y}]+\frac{t^{2}}{2!}[\widehat{X},[\widehat{X}, \widehat{Y}]]+\ldots
$$

This gives the Lie formula for $t=1$.

### 2.13 Weyl and Cambell-Baker-Hausdorff formulas

More difficult is it to derive the Campbell-Baker-Hausdorff formula. In its simplest version it states that

$$
\begin{equation*}
e^{\widehat{X}} e^{\widehat{Y}}=e^{\widehat{X}+\widehat{Y}+\frac{1}{2}[\widehat{X}, \widehat{Y}]} \tag{2.54}
\end{equation*}
$$

which is valid when $[\widehat{X}, \widehat{Y}]=c \widehat{I}$ where $c$ is a constant, i.e. when both $\widehat{X}$ and $\widehat{Y}$ commutes with their commutator. Then it can be proven along the same lines as the Lie formula by now taking the derivative of the function

$$
\widehat{F}(t)=e^{t \widehat{X}} e^{t \widehat{Y}}
$$

which gives

$$
\begin{aligned}
\frac{d \widehat{F}}{d t} & =\widehat{X} e^{t \widehat{X}} e^{t \widehat{Y}}+e^{t \widehat{X}} \widehat{Y} e^{t \widehat{Y}}=\widehat{X} \widehat{F}+e^{t \widehat{X}} \widehat{Y} e^{-t \widehat{X}} e^{t \widehat{X}} e^{t \widehat{Y}} \\
& =\left(\widehat{X}+e^{t \widehat{X}} \widehat{Y} e^{-t \widehat{X}}\right) \widehat{F}=(\widehat{X}+\widehat{Y}+t[\widehat{X}, \widehat{Y}]) \widehat{F}
\end{aligned}
$$

where we have used the Lie formula in the last step. It is now straightforward to show that the solution of this differential equation is

$$
\widehat{F}(t)=e^{t(\widehat{X}+\widehat{Y})+\frac{1}{2} t^{2}[\widehat{X}, \widehat{Y}]}
$$

and the result (2.54) follows by taking here again $t=1$. In the literature this simple result is also often called the Weyl formula.
In the more general case when $\widehat{X}$ and $\widehat{Y}$ do not commute with their commutator $[\widehat{X}, \widehat{Y}]$, we must use brute force in the calculation. Consider

$$
\widehat{Z}=\log \left(e^{\widehat{X}} e^{\widehat{Y}}\right)
$$

and make use of the standard expansion

$$
\begin{equation*}
\log x=(x-1)-\frac{1}{2}(x-1)^{2}+\frac{1}{3}(x-1)^{3}+\ldots \tag{2.55}
\end{equation*}
$$

In the first term of the series we have

$$
\begin{aligned}
e^{\widehat{X}} e^{\widehat{Y}}-1 & =\widehat{X}+\widehat{Y}+\frac{1}{2}\left(\widehat{X}^{2}+2 \widehat{X} \widehat{Y}+\widehat{Y}^{2}\right) \\
& +\frac{1}{6}\left(\widehat{X}^{3}+3 \widehat{X}^{2} \widehat{Y}+3 \widehat{X} \widehat{Y}^{2}+\widehat{Y}^{3}\right)+\ldots
\end{aligned}
$$

where we for simplicity keep only products involving three operators or less. Squaring this, we get

$$
\begin{aligned}
& \left(e^{\widehat{X}} e^{\widehat{Y}}-1\right)^{2}=\widehat{X}^{2}+\widehat{X} \widehat{Y}+\widehat{Y} \widehat{X}+\widehat{Y}^{2} \\
+ & \frac{1}{2}\left(2 \widehat{X}^{3}+3 \widehat{X}^{2} \widehat{Y}+2 \widehat{X} \widehat{Y} \widehat{X}+\widehat{Y} \widehat{X}^{2}+\widehat{Y}^{2} \widehat{X}+2 \widehat{Y} \widehat{X} \widehat{Y}+3 \widehat{X} \widehat{Y}^{2}+2 \widehat{Y}^{3}\right)+\ldots
\end{aligned}
$$

Multiplying these two expansions together, gives then the highest order term we need

$$
\left(e^{\widehat{X}} e^{\widehat{Y}}-1\right)^{3}=\widehat{X}^{3}+\widehat{X}^{2} \widehat{Y}+\widehat{X} \widehat{Y} \widehat{X}+\widehat{Y} \widehat{X}^{2}+\widehat{Y}^{2} \widehat{X}+\widehat{Y} \widehat{X} \widehat{Y}+\widehat{X} \widehat{Y}^{2}+\widehat{Y}^{3}+\ldots
$$

Inserting these results in (2.55) and collecting terms, we find

$$
\begin{align*}
\log \left(e^{\widehat{X}} e^{\widehat{Y}}\right) & =\widehat{X}+\widehat{Y}+\frac{1}{2}(\widehat{X} \widehat{Y}-\widehat{Y} \widehat{X}) \\
& +\frac{1}{12}\left(\widehat{X}^{2} \widehat{Y}+\widehat{Y} \widehat{X}^{2}-2 \widehat{X} \widehat{Y} \widehat{X}+\widehat{Y}^{2} \widehat{X}+\widehat{X} \widehat{Y}^{2}-2 \widehat{Y} \widehat{X} \widehat{Y}\right)+\ldots \\
& \left.=\widehat{X}+\widehat{Y}+\frac{1}{2}[\widehat{X}, \widehat{Y}]+\frac{1}{12}(\widehat{X},[\widehat{X}, \widehat{Y}]]+[\widehat{Y},[\widehat{Y}, \widehat{X}]]\right)+\ldots \tag{2.56}
\end{align*}
$$

which is the Campbell-Baker-Hausdorff formula to third order. Higher order terms can be derived the same way, but have little or zero practical use.

## Chapter 3

## Discrete quantum systems

In the previous chapters we have assumed that the Hilbert spaces of quantum systems have a finite dimension $N$. The corresponding matrix representations can then be labelled by integers $1,2,3, \ldots, N$. Such a system is said to be discrete even in the limit when the dimension $N \rightarrow \infty$. These systems are from a mathematical point usually the simplest to investigate. They also describe many important phenomena in Nature.

### 3.1 Matrix mechanics

In order to solve the abstract Schrödinger equation

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t}|\Psi, t\rangle=\widehat{H}|\Psi, t\rangle \tag{3.1}
\end{equation*}
$$

we need to introduce an orthonormal basis in the $N$-dimensional Hilbert space where the Hamiltonian operator $\widehat{H}$ acts. This can be the complete set of the eigenvectors $\{|n\rangle\}$ of an Hermitian operator $\widehat{A}$ we don't have to specify at this stage. Projecting the equation onto the bra vector $\langle m|$ it becomes

$$
i \hbar \frac{\partial}{\partial t}\langle m \mid \Psi, t\rangle=\langle m| \widehat{H}|\Psi, t\rangle
$$

On the right-hand side we introduce the identity operator so that we can write $\langle m| \widehat{H}|\Psi, t\rangle=$ $\langle m| \widehat{H} \widehat{I}|\Psi, t\rangle$. Expressing now $\widehat{I}$ by the completeness sum, we get

$$
\begin{equation*}
i \hbar \dot{\psi}_{m}=\sum_{n=1}^{N} H_{m n} \psi_{n} \tag{3.2}
\end{equation*}
$$

Here $\psi_{n}(t)=\langle n \mid \Psi, t\rangle$ is the component of $|\Psi, t\rangle$ in the $m$-direction and $H_{m n}=\langle m| \widehat{H}|n\rangle$ gives the matrix elements of $\widehat{H}$ in this basis. Since the state vector

$$
\begin{equation*}
|\Psi, t\rangle=\sum_{n=1}^{N} \psi_{n}(t)|n\rangle \tag{3.3}
\end{equation*}
$$

varies with time, its components will also be time dependent. They are the probability amplitudes for the system to be in state $|n\rangle$ at time $t$ and therefore contain all the quantum mechanical information about the system.

The set of first-order differential equation (3.2) are now equivalent to the abstract Schrödinger equation in this basis. As we have seen previously, we can now represent the state vector by the column matrix $\psi=\left(\psi_{1}, \psi_{2}, \ldots, \psi_{n}\right)^{T}$ and the operator $\widehat{H}$ by the corresponding $N \times N$ matrix $H=\left(H_{m n}\right)$. The equation then takes the matrix form

$$
\begin{equation*}
i \hbar \dot{\psi}=H \psi \tag{3.4}
\end{equation*}
$$

or more specifically

$$
i \hbar\left(\begin{array}{c}
\dot{\psi}_{1}  \tag{3.5}\\
\dot{\psi}_{2} \\
\vdots \\
\dot{\psi}_{N}
\end{array}\right)=\left(\begin{array}{cccc}
H_{11} & H_{12} & \ldots & H_{1 N} \\
H_{21} & H_{22} & \ldots & H_{2 N} \\
\vdots & \vdots & \ddots & \vdots \\
H_{N 1} & H_{N 2} & \ldots & H_{N N}
\end{array}\right)\left(\begin{array}{c}
\psi_{1} \\
\psi_{2} \\
\vdots \\
\psi_{N}
\end{array}\right)
$$

Abstract quantum mechanics in Hilbert space thus becomes matrix mechanics where the probability amplitudes are complex numbers and operators are represented by complex matrices.

The simplest possible system has a one-dimensional Hilbert space and can be in only one state with amplitude $\psi_{1}=\psi$. Since the Hamiltonian matrix then has only one element $H_{11}=E$ which is real, the Schrödinger equation simplifies to $i \hbar \dot{\psi}=E \psi$. It has the elementary solution

$$
\begin{equation*}
\psi(t)=C e^{-i E t / \hbar} \tag{3.6}
\end{equation*}
$$

where $C$ is a integration constant. The probability $|\psi(t)|^{2}=|C|^{2}$ to remain in this state is therefore constant in time as it should be. This is just another way of saying that in such a quantum system nothing happens as time goes by.

### 3.2 The hydrogen molecule

More interesting is the two-state system where the coupled Schrödinger equations (3.2) become

$$
\begin{align*}
& i \hbar \dot{\psi}_{1}=H_{11} \psi_{1}+H_{12} \psi_{2} \\
& i \hbar \dot{\psi}_{2}=H_{21} \psi_{1}+H_{22} \psi_{2} \tag{3.7}
\end{align*}
$$

Since the Hamiltonian is Hermetian $H_{m n}^{*}=H_{n m}$, we find that $H_{11}=E_{1}$ and $H_{22}=E_{2}$ are both real. The non-diagonal elements must be of the form $H_{12}=-A e^{i \delta}=H_{21}^{*}$ with the factor $A$ being real. The phase factor $\delta$ we can choose to be $\delta=0$ since it depends on the relative phase of the two states $|1\rangle$ and $|2\rangle$ which here can be adjusted freely. The Hamiltonian matrix is therefore

$$
H=\left(\begin{array}{rr}
E_{1} & -A  \tag{3.8}\\
-A & E_{2}
\end{array}\right)
$$

and involves only three real parameters. If $A$ had been zero, we immediately see that both $\psi_{1}(t)$ and $\psi_{2}(t)$ varies like the simple exponential in (3.6). If the system thus starts in for instance state $|2\rangle$, it will remain there. But if $A \neq 0$ and the system starts in state $|2\rangle$ so that the amplitude $\psi_{2}(0)=1$, we see that $i \hbar \dot{\psi}_{1}(0)=-A$ and the amplitude $\psi_{1}(t)$ will start to build up from zero. The non-diagonal elements $H_{12}=H_{21}$ therefore cause transitions between the states $|1\rangle$ and $|2\rangle$ and are for this reason called transition matrix elements.

We will first consider the slightly simpler problem where the diagonal elements are equal, $E_{1}=E_{2}=E_{0}$. This is the symmetric case and describes several interesting, physical systems. One is the maser based on the ammonia molecule $N H_{3}$. The three hydrogen atoms form a plane and the nitrogen atom can be above the plane in state $|1\rangle$ or below it in state $|2\rangle$. Needless to say, this is the most coarse-grained description of this complex molecule where we ignore all degrees of freedom except the one decribing the $N$ atom as being above or below the hydrogen plane.


Figure 3.1: Basis states for the 2-state problem of one electron tightly bound to one of two protons.
A mathematically identical system is the ionized hydrogen molecule $H_{2}^{+}$with two protons and one electron. Because of the Coulomb attraction, the electron can be in a state near the first proton, i.e. in state $|1\rangle$ or near the second in state $|2\rangle$ as shown in the Figure 6.1. Again this is a very rough description of the molecule, but it captures the essential physics. We will have this charged molecule in mind in the following.

The first-order differential equations (3.7) now become

$$
\begin{aligned}
i \hbar \dot{\psi}_{1} & =E_{0} \psi_{1}-A \psi_{2} \\
i \hbar \dot{\psi}_{2} & =E_{0} \psi_{2}-A \psi_{1}
\end{aligned}
$$

They are easily solved by taking linear combinations. Adding them, we find

$$
i \hbar \frac{d}{d t}\left(\psi_{1}+\psi_{2}\right)=\left(E_{0}-A\right)\left(\psi_{1}+\psi_{2}\right)
$$

while subtraction gives

$$
i \hbar \frac{d}{d t}\left(\psi_{1}-\psi_{2}\right)=\left(E_{0}+A\right)\left(\psi_{1}-\psi_{2}\right)
$$

The same, direct integration as gave (3.6), now gives the solutions

$$
\begin{aligned}
\psi_{1}(t)+\psi_{2}(t) & =\sqrt{2} C_{-} e^{-i E_{-} t / \hbar} \\
\psi_{1}(t)-\psi_{2}(t) & =\sqrt{2} C_{+} e^{-i E_{+} t / \hbar}
\end{aligned}
$$

where $E_{ \pm}=E_{0} \pm A$ and $C_{ \pm}$are integration constants. The sought-for amplitudes are therefore

$$
\begin{align*}
\psi_{1}(t) & =\sqrt{\frac{1}{2}}\left(C_{-} e^{-i E_{-} t / \hbar}+C_{+} e^{-i E_{+} t / \hbar}\right) \\
\psi_{2}(t) & =\sqrt{\frac{1}{2}}\left(C_{-} e^{-i E_{-} t / \hbar}-C_{+} e^{-i E_{+} t / \hbar}\right) \tag{3.9}
\end{align*}
$$

The probabilities $P_{i}=\left|\psi_{i}\right|^{2}$ for the system to be in state $|i\rangle$ are therefore in general functions of time, depending on the value of the integration constants $C_{ \pm}$. Normally these are determined by the initial conditions.
To be more concrete, let us again think of the $H_{2}^{+}$molecule. At time $t=0$ let us assume that the electron is found to be at proton 1, i.e. the electron is in the state $|1\rangle$ with certainty. Thus $C_{-}=C_{+}=\sqrt{1 / 2}$, ignoring again an irrelevant phase, and the probability amplitudes become

$$
\begin{equation*}
\psi_{1}(t)=e^{-i E_{0} t / \hbar} \cos A t / \hbar, \quad \psi_{2}(t)=i e^{-i E_{0} t / \hbar} \sin A t / \hbar \tag{3.10}
\end{equation*}
$$

From the corresponding probabilities $P_{1}(t)=\cos ^{2} A t / \hbar$ and $P_{2}(t)=\sin ^{2} A t / \hbar$ we see that the electron oscillates between the two protons with frequency $\omega=A / \hbar$. For the similar $\mathrm{NH}_{3}$ molecule this would be the maser frequency. Notice that $P_{1}+P_{2}=1$ at all times as it should for physical probabilities.

But we can also choose other initial conditions so that their is no time dependence in the position of the electron. For both the cases $\left(C_{-}=1, C_{+}=0\right)$ and ( $C_{-}=0, C_{+}=1$ ) we find that the probabilities $P_{1}=P_{2}=1 / 2$ are constant in time. There is the same probability to find the electron at either proton. It is then in one of two posssible stationary states with well-defined energies $E_{ \pm}=E_{0} \pm A$. Needless to say, these are the eigenvalues of the Hamiltonian for the system.
Had we tried to calculate the parameters $E_{0}$ and $A$ from the given Coulomb interactions between the electron and the protons, we would have found that they depended on the distance $r$ between the two protons giving rise to an effective potential $V(r)$. Plotting the stationary energies as in Fig. 6.2, we would then find that the lowest energy $E_{-}=E_{0}-A$ has a minimum at for a certain separation $r_{0}$. This will then be the stable configuration for the molecule and we say that the electron is in an bonding orbital. In the other state corresponding to the energy $E_{+}=E_{0}+A$, there is no stable minimum and the electron is in an anti-bonding orbital where the molecule is unstable.

### 3.3 The general two-state system

The general two-state system is given by the Hamiltonian matrix (3.8). Instead of calculating again the time evolution of the two probability amplitudes, we will now first find the corresponding energy eigenvalues. As explained in the previous chapter, the corresponding eigenvectors are stationary states and form an othonormal basis which is natural to call an energy basis. The energy eigenvalues follow from the matrix equation

$$
\left(\begin{array}{rr}
E_{1} & -A \\
-A & E_{2}
\end{array}\right)\binom{u_{1}}{u_{2}}=E\binom{u_{1}}{u_{2}}
$$

or $\left(E_{1}-E\right) u_{1}=A u_{2}$ and $\left(E_{2}-E\right) u_{2}=A u_{1}$ when written out. These two equations can be solved for the amplitudes $u_{1}$ and $u_{2}$ when

$$
\left|\begin{array}{cc}
E_{1}-E & -A \\
-A & E_{2}-E
\end{array}\right|=0
$$



Figure 3.2: Effective potentials for an electron in the ionized hydrogen molecule as function of the proton separation $r$.

This determinant gives the second order equation $\left(E-E_{1}\right)\left(E-E_{2}\right)-A^{2}=0$ with the two roots

$$
\begin{equation*}
E_{ \pm}=\frac{E_{1}+E_{2}}{2} \pm \sqrt{A^{2}+\left(\frac{E_{1}-E_{2}}{2}\right)^{2}} \tag{3.11}
\end{equation*}
$$

The eigenvector corresponding to $E_{+}$will then have amplitudes satisfying the equation $\left(E_{1}-E_{+}\right) u_{1}=A u_{2}$. Defining $\Delta=\left(E_{1}-E_{2}\right) / 2$, we thus have $u_{1} / u_{2}=A /(\Delta-$ $\sqrt{A^{2}+\Delta^{2}}$ ). This can be simplified by introducing the mixing angle

$$
\begin{equation*}
\tan 2 \theta=\frac{2 A}{E_{1}-E_{2}}=\frac{A}{\Delta} \tag{3.12}
\end{equation*}
$$

which gives $u_{1} / u_{2}=\sin 2 \theta /(\cos 2 \theta-1)=-\cot \theta$. We thus have for the normalized eigenvector

$$
\begin{equation*}
\left|E_{+}\right\rangle=\cos \theta|1\rangle-\sin \theta|2\rangle \tag{3.13}
\end{equation*}
$$

Notice that the second equation $\left(E_{2}-E_{+}\right) u_{2}=A u_{1}$ is now automatically satisfied. Similarly, for the eigenvalue $E_{-}$the amplitudes follow from $\left(E_{1}-E_{-}\right) u_{1}=A u_{2}$ which now gives $u_{1} / u_{2}=\tan \theta$ and therefore

$$
\begin{equation*}
\left|E_{-}\right\rangle=\sin \theta|1\rangle+\cos \theta|2\rangle \tag{3.14}
\end{equation*}
$$

It is seen to be orthogonal to the first eigenvector as it should be.

If now the system at time $t=0$ is in the state

$$
|\Psi\rangle=C_{+}\left|E_{+}\right\rangle+C_{-}\left|E_{-}\right\rangle
$$

it will at a later time have the state vector

$$
|\Psi, t\rangle=C_{+} e^{-i E_{+} t / \hbar}\left|E_{+}\right\rangle+C_{-} e^{-i E_{-} t / \hbar}\left|E_{-}\right\rangle
$$

The amplitude $\psi_{n}(t)=\langle n \mid \Psi, t\rangle$ to be found in one of the first basis states $|n\rangle$ with $n=(1,2)$ is therefore

$$
\begin{equation*}
\psi_{n}(t)=C_{+} e^{-i E_{+} t / \hbar}\left\langle n \mid E_{+}\right\rangle+C_{-} e^{-i E_{-} t / \hbar}\left\langle n \mid E_{-}\right\rangle \tag{3.15}
\end{equation*}
$$

and the corresponding probability $P_{n}=\left|\psi_{n}(t)\right|^{2}$ will in general be time dependent.
In the special case that the system for example starts out in state $|1\rangle, C_{+}=\cos \theta$ and $C_{-}=\sin \theta$. There will then be a non-zero amplitude $\psi_{2}(t)$ to be found in state $|2\rangle$ at a later time. Since $\left\langle 2 \mid E_{+}\right\rangle=-\sin \theta$ and $\left\langle 2 \mid E_{-}\right\rangle=\cos \theta$,

$$
\begin{aligned}
\psi_{2}(t) & =-\sin \theta \cos \theta\left(e^{-i E_{+} t / \hbar}-e^{-i E_{-} t / \hbar}\right) \\
& =-i \sin 2 \theta e^{i\left(E_{+}+E_{-}\right) t / 2 \hbar} \sin \frac{\left(E_{+}-E_{-}\right) t}{2 \hbar}
\end{aligned}
$$

The probability to make the transition $|1\rangle \rightarrow|2\rangle$ is therefore

$$
\begin{equation*}
P_{1 \rightarrow 2}=\sin ^{2} 2 \theta \sin ^{2} \omega t \tag{3.16}
\end{equation*}
$$

where the oscillating frequency $\omega=\left(E_{+}-E_{-}\right) / 2 \hbar$. The probability to remain in the same state will be $P_{1 \rightarrow 1}=1-P_{1 \rightarrow 2}$ since the sum of these two probabilties must be one.
In the symmetric case when $E_{1}=E_{2} \equiv E_{0}$, the eigenvalues become $E_{ \pm}=E_{0} \pm A$. These correspond to the stationary states we found in the previous note. The mixing angle is seen from (3.12) to become $\theta=\pi / 4$ so that the two eigenstates are

$$
\begin{equation*}
\left|E_{+}\right\rangle=\sqrt{\frac{1}{2}}(|1\rangle-|2\rangle), \quad\left|E_{-}\right\rangle=\sqrt{\frac{1}{2}}(|1\rangle+|2\rangle) \tag{3.17}
\end{equation*}
$$

The transition amplitudes follow now directly from (3.15) and are seen to be exactly the same as we derived previously in (3.9) in the symmetric case describing the ionized hydrogen molecule. This should not come as a surprise.

### 3.4 Neutrino oscillations

In the Standard Model of elementary particles there are three neutrinos $\nu_{e}, \nu_{\mu}$ and $\nu_{\tau}$ occuring primarily in weak interactions. For a long time it was thought that these were massless. But observation during the last ten years have made it clear that this is not so. They seem to have non-zero masses of sizes around $1 \mathrm{eV} / \mathrm{c}^{2}$ or below. More accurately, they are not mass eigenstates of an underlying Hamiltonian, but there can be weak transitions beween them. These effects have been seen in the neutrinos coming from
the fusion processes in the Sun and also the atmosphere where they are produced by incoming cosmic rays. Laboratory experiments have verified these observations.

In order to simplify the discussion, assume that the two neutrinos $\nu_{\mu}$ and $\nu_{\tau}$ can be replaced by a single, effective neutrino $\nu_{a}$. This turns out in many cases to be a very good approximation. We then have a two-state system with $\nu_{e}$ and $\nu_{a}$ playing the roles of the previous basis states labelled 1 and 2 . If a $\nu_{e}$ neutrino is produced at time $t=0$ it will then have a certain probability to be detected as a $\nu_{a}$ neutrino at a later time given by (3.16). Since the states $\left|E_{+}\right\rangle$and $\left|E_{-}\right\rangle$are mass eigenstates, the oscillation frequency $\omega$ is determined by the difference

$$
E_{+}-E_{-}=\sqrt{p^{2} c^{2}+m_{+}^{2} c^{4}}-\sqrt{p^{2} c^{2}+m_{-}^{2} c^{4}}=\frac{c^{4}}{2 E}\left(m_{+}^{2}-m_{-}^{2}\right)
$$

for very high energies $E=p c$. For the frequency we therefore find $\omega=\Delta m^{2} c^{4} / 4 \hbar E$ where $\Delta m^{2}=m_{+}^{2}-m_{-}^{2}$. In the oscillation probability (3.16) one can express the time $t$ the neutrino has moved by the corresponding length $L=c t$. Measuring this in km, the energy $E$ in GeV and the mass difference $\Delta m^{2}$ in $\mathrm{eV}^{2} / c^{4}$, it takes the form

$$
\begin{equation*}
P\left(\nu_{e} \rightarrow \nu_{a}\right)=\sin ^{2} 2 \theta \sin ^{2}\left(1.27 \frac{\Delta m^{2} L}{E}\right) \tag{3.18}
\end{equation*}
$$

This is often called Pontecorvo's formula. After the $\nu_{e}$ neutrino has moved a distance $L \approx E / \Delta m^{2}$ it thus has a maximum probability to be detected as a $\nu_{a}$. The determination of mixing angles like $\theta$ and the neutrino mass differences $\Delta m^{2}$ is today a very important part of research in elementary particle physics and cosmology.

### 3.5 Reflection invariance and symmetry transformations

The two-state systems considered in previous notes became especially simple in the symmetric case when the diagonal elements of the Hamiltonian (3.8) had the same vaule. This is the case of the ionized hydrogen molecule which can be in the two states depicted in Figure 6.3. Since the two protons are identical, we have reflection invariance in a plane normal to the line connecting them and placed halfway between them when there is no electron present to disturb the picture.

However, when an electron is present at for instance proton 1, it will correspond to an electron near proton 2 under this reflection. These two quantum states $|1\rangle$ and $|2\rangle$ are therefore related by an operator $\widehat{P}$ which is called an reflection or parity operator defined by

$$
\begin{equation*}
\widehat{P}|1\rangle=|2\rangle, \quad \widehat{P}|2\rangle=|1\rangle \tag{3.19}
\end{equation*}
$$

Hence it has the composition

$$
\begin{equation*}
\widehat{P}=|1\rangle\langle 2|+|2\rangle\langle 1| \tag{3.20}
\end{equation*}
$$



Figure 3.3: System of two identical sites and reflection plane shown by dashed line.
which shows that it is Hermitean. Squaring it, we find $\widehat{P}^{2}=|1\rangle\langle 1|+|2\rangle\langle 2|=\widehat{I}$ which also follows from $\widehat{P}^{2}|1\rangle=\widehat{P}|2\rangle=|1\rangle$. Thus $\widehat{P}=\widehat{P}^{-1}=\widehat{P}^{\dagger}$ and it is therefore also unitary.
Since the parity operator is Hermitean, it has real eigenvalues $\varepsilon$. Since $\widehat{P}^{2}=\widehat{I}$, they must satisfy $\varepsilon^{2}=1$. Therefore we find the values $\varepsilon= \pm 1$. A state with $\varepsilon=+1$ is said to have positive parity and is symmetric under the interchange of the two protons. Similarly, the state with negative parity $\varepsilon=-1$ is antisymmetric under this interchange.

With the two states $|1\rangle$ and $|2\rangle$ as basis states, the Hamiltonian matrix is

$$
H=\left(\begin{array}{rr}
E_{0} & -A  \tag{3.21}\\
-A & E_{0}
\end{array}\right)
$$

while $\widehat{P}$ is seen from the definition (3.19) to be represented by the matrix

$$
P=\left(\begin{array}{ll}
0 & 1  \tag{3.22}\\
1 & 0
\end{array}\right)
$$

Multiplying these two matrices together, it is easily seen that $H P=P H$. They therefore commute with each other as the corresponding operators also will do,

$$
\begin{equation*}
[\widehat{H}, \widehat{P}]=0 \tag{3.23}
\end{equation*}
$$

In the previous chapter we showed that when two operators commute, one can then find eigenvectors which are common to both. So in our case, the parity eigenvectors will also be energy eigenvectors. This is also consistent with the eigenvectors (3.17) we found in the symmetric case. The lowest energy state $\left|E_{-}\right\rangle$is symmetric and thus have positive parity $\varepsilon=+1$ while the unbound state $\left|E_{+}\right\rangle$is antisymmetric and therefore has negative parity.
Symmetry plays an important role in physics and especially in quantum mechanics. From geometrical or other properties of a system, it follows that it can be in different states which must be physically equivalent. This is like the two states describing the position of the electron on identical protons in the hydrogen molecule. There will then be an operator $\widehat{S}$ which relates one such state $|\Psi\rangle$ to an other state $\left|\Psi^{\prime}\right\rangle$, i.e.

$$
\begin{equation*}
\left|\Psi^{\prime}\right\rangle=\widehat{S}|\Psi\rangle \tag{3.24}
\end{equation*}
$$

Since these two states are equivalent, they must also have the same norm. From $\langle\Psi \mid \Psi\rangle=$ $\left\langle\Psi^{\prime} \mid \Psi^{\prime}\right\rangle$ then follows that

$$
\begin{equation*}
\widehat{S}^{\dagger} \widehat{S}=\widehat{I} \tag{3.25}
\end{equation*}
$$

and the symmetry transformation operator $\widehat{S}$ must be unitary. It is only in the very special case when $\widehat{S}^{2}=\widehat{I}$ that it is also Hermitean. In addition, the equivalence must hold at all times for a stationary system and therefore $\left|\Psi^{\prime}, t\right\rangle=\widehat{S}|\Psi, t\rangle$. Since $|\Psi, t\rangle=$ $\widehat{U}(t)|\Psi\rangle$ where $\widehat{U}(t)=\exp (-i \widehat{H} t / \hbar)$ is the time development operator, we find that the operators $\widehat{S}$ and $\widehat{U}(t)$ commute. As a consequence, it thus follows that

$$
\begin{equation*}
[\widehat{H}, \widehat{S}]=0 \tag{3.26}
\end{equation*}
$$

The corresponding result (3.23) was found above for reflection symmetry by explicitly commuting the two operators. But we now see that it is inherent in the definition of a general symmetry transformation.

### 3.6 Lattice translation invariance

We will now extend the hydrogen molecule we have considered previously, to a giant, one-dimensional crystal consisting of an infinite number of identical atoms or ions equally spaced along a line. This is then a lattice and each atom is called a lattice site. The distance between two nearest neighbouring atoms is called the lattice constant a. Although


Figure 3.4: One-dimensional lattice of identical atoms or ions.
such a one-dimensional lattice can actually be made nowadays, here it primarily will serve as a simple example of a system with translational symmetry. Real crystals are three-dimensional and will be considered briefly at the end of the chapter.
The atoms in the lattice can be labeled by the integers $n=(\ldots,-1,0,1,2,3, \ldots)$ as shown in the figure. The distance of atom numbered $n$ from atom $n=0$ is $x_{n}=n a$. Since the lattice is infinite in size, this labelling is arbitrary in the way we assign the position $n=0$ to one particular atom. This atom could just as well be labelled $n=7$ if all the other atoms also got their label increased by the same amount. Such a relabelling would physically correspond to all the atoms in the lattice be shifted to the left by seven lattice spacings. Nothing in the physics of the system changes by this shift. We say that the system is invariant under such translations to the left or to the right by a whole number times the lattice constant $a$.

We now consider a quantum particle like an electron which exists on this lattice. It can only be localized to one of the lattice sites as for the electron in the ionized hydrogen molecule. When it is found at the site numbered $n$, we say that it is in the position
eigenstate $|n\rangle$. These position eigenstates are again orthonormalized so that $\left\langle n \mid n^{\prime}\right\rangle=$ $\delta_{n n^{\prime}}$ with the completeness relation

$$
\begin{equation*}
\sum_{n=-\infty}^{\infty}|n\rangle\langle n|=\widehat{I} \tag{3.27}
\end{equation*}
$$

All matrix representations will therefore be infinite-dimensional.
In the quantum description we can now introduce a unitary operator $\widehat{T}$ which will be present because of the invariance of the system under the above lattice translations. The simplest translation $\widehat{T}(a)$ is just an elementary shift of one lattice spacing defined by

$$
\begin{equation*}
\widehat{T}(a)|n\rangle=|n+1\rangle \tag{3.28}
\end{equation*}
$$

Since this operator is unitary, $\widehat{T}^{\dagger}=\widehat{T}^{-1}$ where the inverse operator $\widehat{T}^{-1}(a)=\widehat{T}(-a)$ acts like

$$
\begin{equation*}
\widehat{T}^{-1}(a)|n\rangle=|n-1\rangle \tag{3.29}
\end{equation*}
$$

Longer translations can now be built up from products of this fundamental operator. For example, a translation of three lattice spacings will be given by the composite operator $\widehat{T}(3 a)=\widehat{T}(a) \widehat{T}(a) \widehat{T}(a)=\widehat{T}^{3}(a)$ and so on. In the same way as the unit operator can be written as the complete sum in (3.27), we see that the basic translation operator is

$$
\begin{equation*}
\widehat{T}(a)=\sum_{n=-\infty}^{\infty}|n+1\rangle\langle n| \tag{3.30}
\end{equation*}
$$

From this abstract expression we easily verify the different physical properties we have already assigned to it.

In order to study more properties of the translation operator, it is very useful to find its eigenvalues. Since it is not Hermetian, these will in general be complex numbers. But we know that the operator is unitary, so the eigenvalues are just phase factors like $\exp (-i f)$. Denoting the corresponding eigenvectors by $|\psi\rangle$, we thus have

$$
\begin{equation*}
\widehat{T}(a)|\psi\rangle=e^{-i f(a)}|\psi\rangle \tag{3.31}
\end{equation*}
$$

where $f(a)$ is so far some unknown function of the lattice spacing. But it can easily be determined from considering composite translations. For instance, we must have

$$
\widehat{T}(2 a)|\psi\rangle=e^{-i f(2 a)}|\psi\rangle
$$

But now we have $\widehat{T}(2 a)|\psi\rangle=\widehat{T}(a) \widehat{T}(a)|\psi\rangle=\widehat{T}(a) e^{-i f(a)}|\psi\rangle=e^{-2 i f(a)}|\psi\rangle$ and therefore $f(2 a)=2 f(a)$. Extending this reasoning, we similarly find $f(n a)=n f(a)$. This functional equation is now uniquely solved by $f(a)=k a$ where $k$ is some undetermined, real number which is usually called the wave number.

We thus have for the translationally invariant eigenstate

$$
\begin{equation*}
\widehat{T}(a)|\psi\rangle=e^{-i k a}|\psi\rangle \tag{3.32}
\end{equation*}
$$

In the basis defined by the position eigenstates $|n\rangle$, it has components $\psi_{n}=\langle n \mid \psi\rangle$. These complex numbers are the probability amplitudes to find the particle at site $n$ when it is in the state $|\psi\rangle$. Writing

$$
\begin{equation*}
|\psi\rangle=\sum_{n=-\infty}^{\infty} \psi_{n}|n\rangle \tag{3.33}
\end{equation*}
$$

in the eigenvalue equation (3.32), we can write

$$
\begin{aligned}
e^{-i k a}|\psi\rangle & =e^{-i k a} \sum_{n=-\infty}^{\infty} \psi_{n}|n\rangle=\widehat{T}(a)|\psi\rangle=\sum_{n=-\infty}^{\infty} \psi_{n} \widehat{T}(a)|n\rangle \\
& =\sum_{n=-\infty}^{\infty} \psi_{n}|n+1\rangle=\sum_{n=-\infty}^{\infty} \psi_{n-1}|n\rangle
\end{aligned}
$$

where we in the last term have changed summation index $n \rightarrow n-1$. Comparing now the corresponding vector components on both sides of this equation, it is seen that $\psi_{n}=e^{i k a} \psi_{n-1}$. Iterating, we thus have $\psi_{n}=e^{i k a} e^{i k a} \psi_{n-2}$ and so on. This gives $\psi_{n}=e^{i k n a} \psi_{0}$ where $\psi_{0}$ is just some complex normalization constant. We will here make the choice $\psi_{0}=1$ giving the probability amplitudes

$$
\begin{equation*}
\psi_{n}=e^{i k n a} \tag{3.34}
\end{equation*}
$$

In other words, when the particle is in the translationally invariant eigenstate $|\psi\rangle$, it has the same probability $P_{n}=\left|\psi_{n}\right|^{2}=1$ to be in any of all the possible lattice positions. If we really could understand what this physically means, we would also have understood the basic mystery of quantum mechanics. The observant reader will also have noticed that the normalization we have chosen here, implies that ket vector $|\psi\rangle$ has a norm or length which is infinite. That is perhaps unexpected, but does not create any difficulties. In the next chapter we will quantize the motion of a particle moving in a continuous spacetime and find a corresponding state vector also with infinite norm.

The wave function (3.34) describes a harmonic wave along the chain with wave length $\lambda=2 \pi / k$. When this becomes much longer than the lattice separation $a$, many neighbouring sites will have approximately the same wave function or probability amplitude. In this limit we expect the microscopic properties of the lattice not to be important and one should be able to use a continuous description of the chain instead.

### 3.7 Lattice dynamics and energy bands

The motion of a particle on the lattice is given by its Hamiltonion. In the simplest case we assume that it can only jump to the nearest neighbour sites with amplitude $-A$ as for the electron in the hydrogen molecule. A jump to the right would then be given by the operator $-A \widehat{T}$ where $\widehat{T}$ is the translation operator (3.30). Similarly, a jump in the opposite direction is given by $-A \widehat{T}^{\dagger}$. If $E_{0}$ is the energy of the particle sitting on a site and unable to jump, then we can write the full Hamiltonian as

$$
\begin{equation*}
\widehat{H}=E_{0} \widehat{I}-A\left(\widehat{T}+\widehat{T}^{\dagger}\right) \tag{3.35}
\end{equation*}
$$

In the lattice or position basis, it would be represented by an infinite-dimensional matrix with $E_{0}$ along the diagonal and $-A$ along the two bidiagonals.

Even if the energy eigenvalue problem is now given by an infinite-dimensional matrix equation, it can easily be solved. The reason is that the translation operator $\widehat{T}$ is a symmetry operator which commutes with the Hamiltonian. This is easily verified since, $\widehat{H} \widehat{T}=E_{0} \widehat{T}-A\left(\widehat{I}+\widehat{T}^{2}\right)=\widehat{T} \widehat{H}$. We therefore know that the eigenstates $|\psi\rangle$ of $\widehat{T}$ are also energy eigenstates. In order to determine the eigenvalues $E$ from $\widehat{H}|\psi\rangle=E|\psi\rangle$ we insert the expansion (3.33) and get

$$
\begin{aligned}
\widehat{H}|\psi\rangle & =\sum_{n=-\infty}^{\infty} \psi_{n}\left(E_{0}|n\rangle-A|n-1\rangle-A|n+1\rangle\right) \\
& =\sum_{n=-\infty}^{\infty}\left(E_{0} \psi_{n}-A \psi_{n+1}-A \psi_{n-1}\right)|n\rangle
\end{aligned}
$$

after shifting summation indices in the two last terms. Comparing terms on both sides, it then follows that

$$
E \psi_{n}=E_{0} \psi_{n}-A\left(\psi_{n+1}+\psi_{n-1}\right)
$$

But the amplitudes $\psi_{n}$ are already known from (3.34). After cancelling a common factor $e^{i k n a}$, we thus find the eigenvalues

$$
\begin{equation*}
E=E_{0}-A\left(e^{i k a}+e^{-i k a}\right)=E_{0}-2 A \cos k a \tag{3.36}
\end{equation*}
$$

They depend periodically on the wave number $k$ as shown in Fig. 6.5. The energy can take on all values between $E_{0} \pm 2 A$ when $|k|<\pi / a$. This possible range of energies is called an energy band. Maximum energy $E_{0}+2 A$ is found when $k= \pm \pi / a$ which is called the wave numbers for the boundaries of the first Brillouin zone.

At the center of the zone $k=0$ and the energy takes the minimum value $E_{0}-2 A$. Near this bottom of the band the energy is seen to increase $\propto k^{2}$. In the next chapter we will see that the wave number will be proportional to the momentum of the particle moving on a lattice which seems continuous for these low-energy states. The particle energy will therefore be proportional with its squared momentum as for non-relativistic motion. It will therefore seem to have a mass given by the exchange energy $A$ and the lattice constant $a$. This is really a very deep result of quantum mechanics and is used in modern theories of particles where our continuous spacetime is just an approximation to a more fundamental description where space and even time are discrete.

Although this derivation of an energy band is highly idealized, one will find essentially the same result in a more realistic description of a quantum electron moving in a onedimensional crystal. As soon has we have a periodic structure, there will be an energy band. It lies at the root of all modern electronic devices which in general have several such allowed bands. This comes about because the atoms at the lattice sites have in general many eigenstates with energies $E_{0}, E_{1}, \ldots$ for a bound electron. We have here considered only transitions of electrons between atoms all in the ground state with energies $E_{0}$. But the atoms can also be in the first excited state with energy $E_{1}$. Since all the atoms must be identical in order to have translation invariance, they are all in this state. The hopping amplitude between sites with such excited atoms will in general be


Figure 3.5: Energy band within the first Brillouin zone.
different from the hopping amplitude $A$ between ground-state atoms. For each atomic bound state we thus get an energy band as shown in Fig. 6.6 when the atoms are brought together in the crystal. Between the bands there are forbidden energies which an electron is not allowed to have. Depending on the relative position of these bands, the system will be an insulator, semiconductor or a metal. These are phenomena studied in quantum solid state theory and used in todays electronics.

### 3.8 Three-dimensional lattices

Let us for the sake of simplicity consider a cubic crystal with lattice spacing $a$ in all three directions. Every lattice site can then be labelled by a vector $\mathbf{n}=\left(n_{x}, n_{y}, n_{z}\right)$ with components that are positive or negative integers. It can be written as

$$
\begin{equation*}
\mathbf{n}=\sum_{i=1}^{3} n_{i} \mathbf{e}_{i} \tag{3.37}
\end{equation*}
$$

where $\mathbf{e}_{i}$ are orthonormal unit vectors in the three directions. In the corresponding Hilbert space we then have a basis of position states $|\mathbf{n}\rangle=\left|n_{x}\right\rangle\left|n_{y}\right\rangle\left|n_{z}\right\rangle$ which is orthonormal in the ordinary sense that

$$
\begin{equation*}
\left\langle\mathbf{n} \mid \mathbf{n}^{\prime}\right\rangle=\delta_{\mathbf{n n}^{\prime}} \tag{3.38}
\end{equation*}
$$

where $\delta_{\mathbf{n n}^{\prime}}=\delta_{n_{x} n_{x}^{\prime}} \delta_{n_{x} y n_{y}^{\prime}} \delta_{n_{z} n_{z}^{\prime}}$. They also form a complete set so that the unit operator

$$
\begin{equation*}
\widehat{I}=\sum_{\mathbf{n}}|\mathbf{n}\rangle\langle\mathbf{n}| \equiv \sum_{n_{x}, n_{y}, n_{z}=-\infty}^{\infty}\left|n_{x}\right\rangle\left|n_{y}\right\rangle\left|n_{z}\right\rangle\left\langle n_{z}\right|\left\langle n_{y}\right|\left\langle n_{x}\right| \tag{3.39}
\end{equation*}
$$



Figure 3.6: Energy bands in a periodic structure arising from atomic energies.
where we in the following will make use of the compact notation in the first sum.
There is now invariance under independent translations in three different directions. Denoting the operators for a shift of just one lattice spacing in the direction $i=(x, y, z)$ by $\widehat{T}_{i}=\widehat{T}\left(a \mathbf{e}_{i}\right)$, we can build up a finite translation by products of such elementary translations

$$
\begin{equation*}
\widehat{T}(a \mathbf{n})=\widehat{T}_{x}^{n_{x}} \widehat{T}_{y}^{n_{y}} \widehat{T}_{z}^{n_{z}} \tag{3.40}
\end{equation*}
$$

They obviously commute with each other, $\left[\widehat{T}_{i}, \widehat{T}_{j}\right]=0$. The effect of such a general translation operator on the basis vector $\left|\mathbf{n}^{\prime}\right\rangle$ is then

$$
\begin{equation*}
\widehat{T}(a \mathbf{n})\left|\mathbf{n}^{\prime}\right\rangle=\left|\mathbf{n}+\mathbf{n}^{\prime}\right\rangle \tag{3.41}
\end{equation*}
$$

Since all these translation operators are unitary, the inverse $\widehat{T}^{-1}(a \mathbf{n})=\widehat{T}(-a \mathbf{n})$ equals the adjoint operator $\widehat{T}^{\dagger}(a \mathbf{n})$ as for the one-dimensional lattice.

We can now construct translationally invariant states, i.e. eigenstates $|\psi\rangle$ satisfying

$$
\begin{equation*}
\widehat{T}(a \mathbf{n})|\psi\rangle=e^{-i f(a, \mathbf{n})}|\psi\rangle \tag{3.42}
\end{equation*}
$$

where the exponential form of the eigenvalue again follows from unitarity. Since this state must also be an eigenstate of the three separate operators $\widehat{T}_{i}$, we can repeat the arguments we went through in the one-dimensional case for each separate direction. The arbitrary function $f(a, \mathbf{n})$ is then seen to be a sum of the general form $f(a, \mathbf{n})=$ $a\left(k_{x} n_{x}+k_{y} n_{y}+k_{z} n_{z}\right)$ where the numbers $k_{i}$ take real values. Grouping them into the wave number vector $\mathbf{k}=\left(k_{x}, k_{y}, k_{z}\right)$, we then have for the above eigenvalue problem

$$
\begin{equation*}
\widehat{T}(a \mathbf{n})|\psi\rangle=e^{-i a \mathbf{k} \cdot \mathbf{n}}|\psi\rangle \tag{3.43}
\end{equation*}
$$

Expressing the eigenstate by its components $\psi_{\mathbf{n}}=\langle\mathbf{n} \mid \psi\rangle$ in the position basis so that

$$
\begin{equation*}
|\psi\rangle=\sum_{\mathbf{n}} \psi_{\mathbf{n}}|\mathbf{n}\rangle \tag{3.44}
\end{equation*}
$$

we then find

$$
\begin{equation*}
\psi_{\mathbf{n}}=e^{i a \mathbf{k} \cdot \mathbf{n}} \tag{3.45}
\end{equation*}
$$

when we choose the normalization $\psi_{\mathbf{n}=\mathbf{0}}=1$ as in the one-dimensional case. It is easily verified that this result satisfies the eigenvalue equation (3.43)

$$
\begin{aligned}
\widehat{T}(a \mathbf{n})|\psi\rangle & =\widehat{T}(a \mathbf{n}) \sum_{\mathbf{n}^{\prime}} e^{i a \mathbf{k} \cdot \mathbf{n}^{\prime}}\left|\mathbf{n}^{\prime}\right\rangle=\sum_{\mathbf{n}^{\prime}} e^{i a \mathbf{k} \cdot \mathbf{n}^{\prime}}\left|\mathbf{n}+\mathbf{n}^{\prime}\right\rangle \\
& =\sum_{\mathbf{m}} e^{i a \mathbf{k} \cdot(\mathbf{m}-\mathbf{n})}|\mathbf{m}\rangle=e^{-i a \mathbf{k} \cdot \mathbf{n}} \sum_{\mathbf{m}} e^{i a \mathbf{k} \cdot \mathbf{m}}|\mathbf{m}\rangle=e^{-i a \mathbf{k} \cdot \mathbf{n}}|\psi\rangle
\end{aligned}
$$

after shifting summation variables a couple of times. The result (3.45) is the probabilty amplitude for finding an electron at the lattice site $\mathbf{n}$ when it is in the translationally invariant state $|\psi\rangle$ characterized by the wave number vector $\mathbf{k}$.
If we want to find the energy of the electron in such a state, we must know the Hamiltonian. In the simplest case we again assume that an electron can only jump to one of its six nearest neighbours with the amplitude $-A$. The resulting Hamiltonian is then just the three-dimensional version of (3.35),

$$
\begin{equation*}
\widehat{H}=E_{0} \widehat{I}-A \sum_{i=1}^{3}\left(\widehat{T}_{i}+\widehat{T}_{i}^{\dagger}\right) \tag{3.46}
\end{equation*}
$$

The above state $|\psi\rangle$ is then also an energy eigenstate satisfying $\widehat{H}|\psi\rangle=E|\psi\rangle$. With the expansion (3.44) this eigenvalue problem then reduces to

$$
E \psi_{\mathbf{n}}=E_{0} \psi_{\mathbf{n}}-A \sum_{j=1}^{3}\left(\psi_{\mathbf{n}+\mathbf{e}_{j}}+\psi_{\mathbf{n}-\mathbf{e}_{j}}\right)
$$

Since we know the probability amplitudes (3.45), we obtain

$$
\begin{align*}
E & =E_{0}-A \sum_{j=1}^{3}\left(e^{i k_{j} a}+e^{-i k_{j} a}\right)  \tag{3.47}\\
& =E_{0}-2 A\left(\cos k_{x} a+\cos k_{y} a+\cos k_{z} a\right) \tag{3.48}
\end{align*}
$$

after cancelling a common factor on both sides. Again we find an energy band varying between $E_{0}+6 A$ and $E_{0}-6 A$. Near the bottom of the band the energy is again seen to increase like $\propto \mathbf{k}^{2}$ as we discussed in the one-dimensional case. Outside the first Brillouin zone defined by the wave numbers $k_{i}= \pm \pi / a$, the energy eigenvalues repeat themselves periodically.

## Chapter 4

## Schrödinger wave mechanics

In the limit when the spacing $a \rightarrow 0$ the lattice becomes a continuum. A particle which could previously only be found at lattice sites $\mathbf{n}$, can then move freely around. Correspondingly, the probability amplitudes $\psi_{\mathbf{n}}$ for such an observation now become a continuous function of the particle position. It is called the wave function of the particle.

### 4.1 Coordinate basis

Introducing the coordinate distance $x=n a$ for the site labelled by $n$ on a one-dimensional lattice, we can denote the discrete basis vectors $|n\rangle$ in the previous chapter by $\left|x_{n}\right\rangle$. The normalization $\left\langle n \mid n^{\prime}\right\rangle=\delta_{n n^{\prime}}$ can then just as well be written as $\left\langle x_{n} \mid x_{n^{\prime}}\right\rangle=\delta_{n n^{\prime}}$. Similarly, the completeness relation (3.27) becomes

$$
\widehat{I}=\sum_{n=-\infty}^{\infty}\left|x_{n}\right\rangle\left\langle x_{n}\right|
$$

When now the coordinate $x_{n}$ takes on all continuous values $x$, the natural generalization of this relation is

$$
\begin{equation*}
\widehat{I}=\int_{-\infty}^{\infty} d x|x\rangle\langle x| \tag{4.1}
\end{equation*}
$$

Here we must now consider $|x\rangle$ as the eigenvector for the position operator $\widehat{x}$ belonging to the eigenvalue $x$,

$$
\begin{equation*}
\widehat{x}|x\rangle=x|x\rangle \tag{4.2}
\end{equation*}
$$

and correspondingly $\widehat{x}\left|x^{\prime}\right\rangle=x^{\prime}\left|x^{\prime}\right\rangle$. We can then write

$$
\left|x^{\prime}\right\rangle=\widehat{I}\left|x^{\prime}\right\rangle=\int_{-\infty}^{\infty} d x|x\rangle\left\langle x \mid x^{\prime}\right\rangle
$$

For consistency, we thus find that the matrix element $\left\langle x \mid x^{\prime}\right\rangle$ which is a function of the two coordinates $x$ and $x^{\prime}$, can be expressed as

$$
\begin{equation*}
\left\langle x \mid x^{\prime}\right\rangle=\delta\left(x-x^{\prime}\right) \tag{4.3}
\end{equation*}
$$

when we make use of the Dirac $\delta$-function. For any smooth function $f(x)$, it is in general defined by the integral

$$
\begin{equation*}
\int_{-\infty}^{\infty} d x f(x) \delta(x-a)=f(a) \tag{4.4}
\end{equation*}
$$

It contributes only in the point $x=a$ and gives zero otherwise. We will say a bit more about it in the following.

We thus see that in this basis Hilbert space not only has infinite dimension, but all matrix representations are labelled by continuous indices. From (4.3) we also see that the basis vectors have infinite norm. But luckily all these apparent complications change very little of the mathematical structure we have previously bult up. The main difference is that all summation over indices are now replaced by integrations.

### 4.2 The Dirac $\delta$-function

From the definition (4.4) we see that the Dirac $\delta$-function $\delta(x)$ which is zero when $x \neq 0$, must satisfy the integral requirement

$$
\begin{equation*}
\int_{-\infty}^{\infty} d x \delta(x)=1 \tag{4.5}
\end{equation*}
$$

In some way it must then be infinite at the point $x=0$. One possibility would therefore be to define it by the limiting process

$$
\delta(x)=\lim _{\varepsilon \rightarrow 0} \begin{cases}1 / \varepsilon, & |x| \leq \varepsilon / 2 \\ 0, & |x|>\varepsilon / 2\end{cases}
$$

But for any small, but non-zero value of $\varepsilon$ the expression on the right-hand side is very discontinuous and mathematically repellent. A much more useful definition is to take the equivalent expression

$$
\begin{equation*}
\delta(x)=\frac{1}{\pi} \lim _{\varepsilon \rightarrow 0} \frac{\varepsilon}{x^{2}+\varepsilon^{2}} \tag{4.6}
\end{equation*}
$$

where now the right-hand side is a smooth function for any non-zero value of $\varepsilon$. The factor $1 / \pi$ in front is needed to satisfy the normalization (4.5) using the standard integral

$$
\int_{-\infty}^{\infty} \frac{d x}{1+x^{2}}=\pi
$$

What makes this definition in particular so attractive, is that we can now rewrite the above limit formula as the integral

$$
\delta(x)=\lim _{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} \frac{d k}{2 \pi} e^{i k x-\varepsilon|k|}=\frac{1}{2 \pi} \lim _{\varepsilon \rightarrow 0}\left(\frac{1}{\varepsilon-i x}+\frac{1}{\varepsilon+i x}\right)
$$

which is what we want. Taking now the limit $\varepsilon \rightarrow 0$ in the integral, we have the most useful reresentation of the $\delta$-function

$$
\begin{equation*}
\delta(x)=\int_{-\infty}^{\infty} \frac{d k}{2 \pi} e^{i k x} \tag{4.7}
\end{equation*}
$$

As it is written now, it is no longer strictly convergent at the upper and lower limits. However, whenever we meet it, we define it by this limiting procedure and thus end up with the Dirac $\delta$-function.

One can also make the integral (4.7) well defined by making the lower and upper limits finite by introducing a cut-off and then later taking the cut-off to infinity,

$$
\begin{align*}
\delta(x) & =\lim _{K \rightarrow \infty} \int_{-K}^{K} \frac{d k}{2 \pi} e^{i k x}=\left.\frac{1}{2 \pi i x} \lim _{K \rightarrow \infty} e^{i k x}\right|_{-K} ^{K} \\
& =\lim _{K \rightarrow \infty} \frac{\sin K x}{\pi x} \tag{4.8}
\end{align*}
$$

There are obviously many other limiting expressions for this very special and important function.

The $\delta$-function is obviously an even function, $\delta(x)=\delta(-x)$. Also, from the definition

$$
\int_{-\infty}^{\infty} d x f(x) \delta(x)=f(0)
$$

where $f(x)$ is a continuous function, follows directly that $x \delta(x)=0$ when we make the special choice $f(x)=x$. It has the useful property

$$
\begin{equation*}
\delta(a x)=\frac{1}{|a|} \delta(x) \tag{4.9}
\end{equation*}
$$

which is seen from a shift $x \rightarrow a x$ of the integration variable in the definition. Although highly singular, we can also introduce the derivative $\delta^{\prime}(x)$. It satisfies

$$
\begin{equation*}
x \delta^{\prime}(x)=-\delta(x) \tag{4.10}
\end{equation*}
$$

as is seen from the integral

$$
\begin{aligned}
\int_{-\infty}^{\infty} d x f(x) x \delta^{\prime}(x) & =\left.f(x) x \delta(x)\right|_{-\infty} ^{\infty} \\
& -\int_{-\infty}^{\infty} d x \delta(x)\left[f(x)+x f^{\prime}(x)\right]=-f(0)
\end{aligned}
$$

since the term with $x \delta(x)$ in the last integral gives zero.

### 4.3 Continuous translations and momentum

In the previous chapter we introduced the translation operator $\widehat{T}(n a)$ that moved the lattice $n$ lattice sites. In terms of the position eigenvectors, their action could be summed up in $\widehat{T}(n a)\left|n^{\prime}\right\rangle=\left|n+n^{\prime}\right\rangle$. When we take the limit $a \rightarrow 0$, we end up with continuous translations given by the corresponding operators $\widehat{T}(x)$ with the equivalent effect

$$
\begin{equation*}
\widehat{T}(x)\left|x^{\prime}\right\rangle=\left|x+x^{\prime}\right\rangle \tag{4.11}
\end{equation*}
$$

These are obviously also unitary in the sense that $\widehat{T}^{\dagger}(x)=\widehat{T}^{-1}(x)=\widehat{T}(-x)$ where $\widehat{T}(0)=\widehat{I}$ is the unit operator.

The eigenstates $|\psi\rangle$ of these continuous translation operators will play a central role in this chapter. From the discrete derivation summed up in (3.32), they will obey the eigenvalue equation

$$
\begin{equation*}
\widehat{T}(x)\left|\psi_{k}\right\rangle=e^{-i k x}\left|\psi_{k}\right\rangle \tag{4.12}
\end{equation*}
$$

where we have labelled them by the corresponding wave number $k$. They can be expanded in a complete set of coordinate eigenstates, i.e.

$$
\begin{equation*}
\left|\psi_{k}\right\rangle=\int_{-\infty}^{\infty} d x \psi_{k}(x)|x\rangle \tag{4.13}
\end{equation*}
$$

where $\psi_{k}(x)=\left\langle x \mid \psi_{k}\right\rangle$ is the amplitude for finding a particle in the state $\left|\psi_{k}\right\rangle$ at position $x$. We call this the wave function for the particle. Applying now the operator $\widehat{T}(x)$ on this equation and using (4.12), it follows that

$$
\psi_{k}(x)=e^{i k x^{\prime}} \psi_{k}\left(x-x^{\prime}\right)
$$

With the same normalization $\psi_{k}(0)=1$ as in the previous chapter, we then have the explicit result

$$
\begin{equation*}
\psi_{k}(x)=e^{i k x} \tag{4.14}
\end{equation*}
$$

for the wave function. It is a plane wave with wave number $k$ which was also the reason for originally giving it that name.
While the evolution operator $\widehat{U}(t)=\exp (-i \widehat{H} t / \hbar)$ translated a state from an intial time $t=0$ to the corresponding state at a later time $t$, the translation operator $\widehat{T}(x)$ is the similar operator for translations in space, here along the $x$-axis. Both are unitary and both equal the unit operator in the limit when their arguments go to zero. We saw in Chapter 2 how the Hamiltonian is the generator of infinitesemal translations in time. The corresponding generator for infinitesemal translations in space is the momentum operator $\widehat{p}$ so that

$$
\begin{equation*}
\widehat{T}(x)=e^{-i \widehat{p x} x / \hbar} \tag{4.15}
\end{equation*}
$$

The requirement of unitarity now implies that $\widehat{p}$ is a Hermitian operator. It thus has real eigenvalues which follows from (4.12) which now can be written as

$$
e^{-i \widehat{p x} / \hbar}\left|\psi_{k}\right\rangle=e^{-i k x}\left|\psi_{k}\right\rangle
$$

The translationally invariant state $\left|\psi_{k}\right\rangle$ therefore must satisfy

$$
\widehat{p}\left|\psi_{k}\right\rangle=\hbar k\left|\psi_{k}\right\rangle
$$

and is thus an eigenstate of the momentum operator with eigenvalue $p=\hbar k$. Introducing the more convenient notation $\left|\psi_{k}\right\rangle \rightarrow|p\rangle$ for these momentum eigenstates, we then have simply

$$
\begin{equation*}
\widehat{p}|p\rangle=p|p\rangle \tag{4.16}
\end{equation*}
$$

We already know that a particle in this state have the wavefunction

$$
\begin{equation*}
\psi_{p}(x)=\langle x \mid p\rangle=e^{i p x / \hbar} \tag{4.17}
\end{equation*}
$$

It follows in an equivalent way from the equation

$$
\widehat{T}(x)|p\rangle=e^{-i p x / \hbar}|p\rangle
$$

and projecting it on the bra vector $\langle x|$ making use of $\langle x| \widehat{T}(x)=\langle 0|$. We see again that the probability to find the particle at a given position is constant and the same for all values of the coordinate $x$.

Notice that this wave function (4.17) has a normalization which is different from the usual one in standard textbooks. This will also show up in the normalization of the momentum eigenstates (4.16). We see that from inserting a complete set of coordinate eigenstates in the matrix element $\left\langle p \mid p^{\prime}\right\rangle$ which gives

$$
\left\langle p \mid p^{\prime}\right\rangle=\int_{-\infty}^{\infty} d x\langle p \mid x\rangle\left\langle x \mid p^{\prime}\right\rangle
$$

The matrix elements under the integral sign are just the wave function (4.17) and its complex conjugate so that

$$
\left\langle p \mid p^{\prime}\right\rangle=\int_{-\infty}^{\infty} d x e^{-i\left(p-p^{\prime}\right) x / \hbar}=2 \pi \hbar \delta\left(p-p^{\prime}\right)
$$

when we make use of the $\delta$-function integral (4.7). Except for the factor $2 \pi \hbar$ on the righthand side, this is the same and infinite norm we also had for the position eigenstates in (4.3). The same factor also shows up in the corresponding completeness integral

$$
\begin{equation*}
\int_{-\infty}^{\infty} \frac{d p}{2 \pi \hbar}|p\rangle\langle p|=\widehat{I} \tag{4.18}
\end{equation*}
$$

as is easily verified by inserting it into the coordinate basis norm $\left\langle x \mid x^{\prime}\right\rangle$ and comparing with the defining normalization (4.3). As a systematic rule, which is easy to remember, we always have $2 \pi \hbar \delta(p)$ in momentum $\delta$-functions and correspondingly $\int d p / 2 \pi \hbar$ in momentum integrations. The last case is equivalent to what is found in statistical mechanics when one sums up states in momentum or phase space.

### 4.4 Operators in the coordinate representation

For a particle moving in continuous space we now have two complete sets of basis vectors we can use to provide a matrix representation of states and operators. Most useful is the coordinate representation based on the position eigenvectors $|x\rangle$. An arbitraty state $|\psi\rangle$ has the components

$$
\begin{equation*}
\psi(x)=\langle x \mid \psi\rangle \tag{4.19}
\end{equation*}
$$

in this basis. This is just the corresponding wavefunction. Making use of the completeness relation (4.1), we thus see that the inner product of two such states is given by the integral

$$
\begin{equation*}
\left\langle\psi_{i} \mid \psi_{j}\right\rangle=\int_{-\infty}^{\infty} d x\left\langle\psi_{i} \mid x\right\rangle\left\langle x \mid \psi_{j}\right\rangle=\int_{-\infty}^{\infty} d x \psi_{i}^{*}(x) \psi_{j}(x) \tag{4.20}
\end{equation*}
$$

It is finite for wave functions localized to a finite region of space. We have already seen that this is not the case for the momentum eigenstates.

Operators will be represented by continuous matrices in this basis. The simplest example is

$$
\begin{equation*}
\langle x| \widehat{x}\left|x^{\prime}\right\rangle=x \delta\left(x-x^{\prime}\right) \tag{4.21}
\end{equation*}
$$

since the basis vectors are eigenstates of the position operator $\widehat{x}$. For the same reason we have $\langle x| \widehat{x}^{n}|\psi\rangle=x^{n}\langle x \mid \psi\rangle$ where now $|\psi\rangle$ is a general state. Using now the Taylor expansion of general operator function $f(\widehat{x})$, it thus follows that

$$
\begin{equation*}
\langle x| f(\widehat{x})|\psi\rangle=f(x)\langle x \mid \psi\rangle \tag{4.22}
\end{equation*}
$$

All operators involving the position operator alone are seen to be diagonal in this basis. This is not the case with the momentum operator. But the non-diagonal matrix element $\langle x| \hat{p}|p\rangle$ have an eigenstate on the right and therefore simplifies to

$$
\langle x| \widehat{p}|p\rangle=p\langle x \mid p\rangle=p e^{i p x / \hbar}=-i \hbar \frac{\partial}{\partial x}\langle x \mid p\rangle
$$

For an arbitrary state

$$
|\psi\rangle=\int_{-\infty}^{\infty} \frac{d p}{2 \pi \hbar}|p\rangle\langle p \mid \psi\rangle
$$

we then get the more general matrix element

$$
\begin{equation*}
\langle x| \widehat{p}|\psi\rangle=-i \hbar \frac{\partial}{\partial x} \int_{-\infty}^{\infty} \frac{d p}{2 \pi \hbar}\langle x \mid p\rangle\langle p \mid \psi\rangle=-i \hbar \partial_{x}\langle x \mid \psi\rangle \tag{4.23}
\end{equation*}
$$

The coordinate matrix elements of the momentum operator is therefore

$$
\begin{equation*}
\langle x| \widehat{p}\left|x^{\prime}\right\rangle=-i \hbar \partial_{x}\left\langle x \mid x^{\prime}\right\rangle=-i \hbar \delta^{\prime}\left(x-x^{\prime}\right) \tag{4.24}
\end{equation*}
$$

Similarly, more complex matrix elements follow from

$$
\langle x| \widehat{p}^{n}|\psi\rangle=\left(-i \hbar \partial_{x}\right)^{n}\langle x \mid \psi\rangle
$$

by the same reasoning. As consequence, one then has for the matrix element of a function of the same operator

$$
\begin{equation*}
\langle x| g(\widehat{p})|\psi\rangle=g\left(-i \hbar \partial_{x}\right)\langle x \mid \psi\rangle \tag{4.25}
\end{equation*}
$$

which corresponds to (4.22). When both operators $\widehat{x}$ and $\widehat{p}$ are present in an operator function $F(\widehat{x}, \widehat{p})$, we get the combined result

$$
\begin{equation*}
\langle x| F(\widehat{x}, \widehat{p})|\psi\rangle=F\left(x,-i \hbar \partial_{x}\right) \psi(x) \tag{4.26}
\end{equation*}
$$

Since the two operators don't commute with each other, their mutual ordering is here important. We thus have that in the coordinate representation we can replace the position operator $\widehat{x}$ simply by the eigenvalue $x$ while for the momentum operator we must make the replacement $\widehat{p} \rightarrow-i \hbar \partial_{x}$. This is known from more elementary courses.

We are now in the position to actually calculate the commutator between $\widehat{x}$ and $\hat{p}$. First we have that

$$
\langle x| \widehat{x} \widehat{p}|\psi\rangle=x\langle x| \widehat{p}|\psi\rangle=-i \hbar x \partial_{x}\langle x \mid \psi\rangle
$$

since $\langle x|$ is an eigenstate of $\widehat{x}$. Then, since $\langle x| \widehat{p} \widehat{x}|\psi\rangle=\langle x| \widehat{p}\left|\psi^{\prime}\right\rangle$ with $\left|\psi^{\prime}\right\rangle=\widehat{x}|\psi\rangle$, one similarly obtains

$$
\langle x| \widehat{p} \widehat{x}|\psi\rangle=-i \hbar \partial_{x} \psi^{\prime}(x)=-i \hbar \partial_{x}(x \psi)=-i \hbar\left(\psi+x \partial_{x} \psi\right)
$$

Subtracting these two partial results, it follows that

$$
\langle x|[\widehat{x}, \widehat{p}]|\psi\rangle=i \hbar\langle x \mid \psi\rangle
$$

Since this is now valid for any state $\psi$, we must have the general result

$$
\begin{equation*}
[\widehat{x}, \widehat{p}]=i \hbar \tag{4.27}
\end{equation*}
$$

This is the basic or canonical commutator and was derived by Heisenberg in his very first paper on quantum mechanics. In our treatment it follows from the underlying assumptions previously made while others impose it as the basic quantum postulate. On the right-hand side of (4.27) and in other similar situations, we have simply dropped the unit operator since it just gives a factor one when acting on a state.

From the above canonical commutator we can now evaluate other, related commutators. As an example, consider $\left[\widehat{x}, \widehat{p}^{2}\right]$. Using the ordinary rules for commutators, then follows that

$$
\left[\widehat{x}, \widehat{p}^{2}\right]=\widehat{p}[\widehat{x}, \widehat{p}]+[\widehat{x}, \widehat{p}] \widehat{p}=2 i \hbar \widehat{p}
$$

This is easily generalized to

$$
\begin{equation*}
\left[\widehat{x}, \hat{p}^{n}\right]=i \hbar n \widehat{p}^{n-1} \tag{4.28}
\end{equation*}
$$

which is thus true for $n=2$. Now

$$
\left[\widehat{x}, \widehat{p}^{n}\right]=\widehat{p}\left[\widehat{x}, \widehat{p}^{n-1}\right]+i \hbar \widehat{p}^{n-1}
$$

and the result follows by induction. We see that the operator $\widehat{x}$ again acts like the differential operator $i \hbar \partial_{p}$. Therefore we will also have more generally

$$
\begin{equation*}
[\widehat{x}, g(\widehat{p})]=\left.i \hbar \partial_{p} g(p)\right|_{p \rightarrow \widehat{p}} \tag{4.29}
\end{equation*}
$$

An identical derivation will then also give the similar commutators

$$
\begin{equation*}
\left[\widehat{x}^{n}, \widehat{p}\right]=i \hbar n \widehat{x}^{n-1} \tag{4.30}
\end{equation*}
$$

and

$$
\begin{equation*}
[f(\widehat{x}), \widehat{p}]=\left.i \hbar \partial_{x} f(x)\right|_{x \rightarrow \widehat{x}} \tag{4.31}
\end{equation*}
$$

which shows that the momentum operator $\widehat{p}$ acts like $-i \hbar \partial_{x}$ as expected.

### 4.5 The momentum representation

We can now go through a similar derivation using instead the momentum basis. Projecting the state $\psi$ upon these basis vectors, we get the wave function in momentum space

$$
\begin{equation*}
\tilde{\psi}(p)=\langle p \mid \psi\rangle \tag{4.32}
\end{equation*}
$$

It is directly related to the corresponding wavefunction $\psi(x)$ in coordinate space as is seen by inserting a complete set of coordinate basis vectors,

$$
\begin{equation*}
\tilde{\psi}(p)=\int_{-\infty}^{\infty} d x\langle p \mid x\rangle\langle x \mid \psi\rangle=\int_{-\infty}^{\infty} d x \psi(x) e^{-i p x / \hbar} \tag{4.33}
\end{equation*}
$$

This this just an ordinary Fourier integral. The inverse transformation

$$
\begin{equation*}
\psi(x)=\int_{-\infty}^{\infty} \frac{d p}{2 \pi \hbar} \tilde{\psi}(p) e^{i p x / \hbar} \tag{4.34}
\end{equation*}
$$

follows by inserting similarly a complete set of momentum eigenstates in the wavefunction $\psi(x)=\langle x \mid \psi\rangle$.

Repeating now what we did in coordinate space for matrix elements of operators, we first get

$$
\langle p| \widehat{x}|x\rangle=x\langle p \mid x\rangle=p e^{-i p x / \hbar}=i \hbar \partial_{p}\langle p \mid x\rangle
$$

and therefore

$$
\langle p| \widehat{x}|\psi\rangle=i \hbar \partial_{p}\langle p \mid \psi\rangle
$$

in analogy with (4.23). A few more steps then give the momentum version of (4.26),

$$
\begin{equation*}
\langle p| F(\widehat{x}, \widehat{p})|\psi\rangle=F\left(i \hbar \partial_{p}, p\right) \tilde{\psi}(p) \tag{4.35}
\end{equation*}
$$

The net result is that if we want to use the momentum basis, the position operator must be replaced by $\widehat{x} \rightarrow i \hbar \partial_{p}$.

### 4.6 Three dimensions

Most physical systems of particles are in three spatial dimensions. The generalization of our previous, one-dimensional results is however straightforward. Positions of particles are given by vectors

$$
\begin{equation*}
\mathbf{x}=\sum_{i=1}^{3} x_{i} \mathbf{e}_{i} \tag{4.36}
\end{equation*}
$$

with orthonormalized basis vectors $\mathbf{e}_{i} \cdot \mathbf{e}_{j}=\delta_{i j}$. These are eigenvalues of the vector position operator

$$
\begin{equation*}
\widehat{\mathbf{x}}=\sum_{i=1}^{3} \widehat{x}_{i} \mathbf{e}_{i} \tag{4.37}
\end{equation*}
$$

with eigenstates $|\mathbf{x}\rangle=|x\rangle|y\rangle|z\rangle$ satisfying the eigenvalue equation

$$
\begin{equation*}
\widehat{\mathrm{x}}|\mathrm{x}\rangle=\mathrm{x}|\mathrm{x}\rangle \tag{4.38}
\end{equation*}
$$

Notice that this actually represents three equations, one in each direction. The normalization in each direction will be the same as in the one-dimensional case so that

$$
\begin{equation*}
\left\langle\mathbf{x} \mid \mathbf{x}^{\prime}\right\rangle=\delta\left(\mathbf{x}-\mathbf{x}^{\prime}\right) \tag{4.39}
\end{equation*}
$$

where we define $\delta\left(\mathbf{x}-\mathbf{x}^{\prime}\right)=\delta\left(x-x^{\prime}\right) \delta\left(y-y^{\prime}\right) \delta\left(z-z^{\prime}\right)$. Thus we can write the completeness relation as

$$
\begin{equation*}
\int d^{3} x|\mathbf{x}\rangle\langle\mathbf{x}|=\widehat{I} \tag{4.40}
\end{equation*}
$$

Needless to say, this is consistent with the normalization (4.39) as is seen by inserting this unit operator into the matrix element on the left.

There will ow be three independent translation operators, one in each direction. The combined operator $\widehat{T}(\mathbf{x})=\widehat{T}(x) \widehat{T}(y) \widehat{T}(z)$ can be defined by

$$
\begin{equation*}
\widehat{T}(\mathbf{x})\left|\mathbf{x}^{\prime}\right\rangle=\left|\mathbf{x}+\mathbf{x}^{\prime}\right\rangle \tag{4.41}
\end{equation*}
$$

and has the general form

$$
\begin{equation*}
\widehat{T}(\mathbf{x})=e^{-i \widehat{\mathbf{p}} \cdot \mathbf{x} / \hbar} \tag{4.42}
\end{equation*}
$$

where the momentum vector operator has three components $\widehat{\mathbf{p}}=\left(\widehat{p}_{x}, \widehat{p}_{y}, \widehat{p}_{z}\right)$ that commute with each other. Its eigenstates $|\mathbf{p}\rangle$ have the normalization

$$
\begin{equation*}
\left\langle\mathbf{p} \mid \mathbf{p}^{\prime}\right\rangle=(2 \pi \hbar)^{3} \delta\left(\mathbf{p}-\mathbf{p}^{\prime}\right) \tag{4.43}
\end{equation*}
$$

and completeness relation

$$
\begin{equation*}
\int \frac{d^{3} p}{(2 \pi \hbar)^{3}}|\mathbf{p}\rangle\langle\mathbf{p}|=\widehat{I} \tag{4.44}
\end{equation*}
$$

also follow from what we have done before. These momentum eigenstates have the coordinate wavefunctions

$$
\begin{equation*}
\psi_{\mathbf{p}}(\mathbf{x})=\langle\mathbf{x} \mid \mathbf{p}\rangle=e^{i \mathbf{p} \cdot \mathbf{x} / \hbar} \tag{4.45}
\end{equation*}
$$

which are obviously three-dimensional plane waves. A general wave function in the coordinate representation

$$
\begin{equation*}
\psi(\mathbf{x})=\langle\mathbf{x} \mid \psi\rangle \tag{4.46}
\end{equation*}
$$

can be expressed in terms of the momentum wavefunction $\tilde{\psi}(\mathbf{p})=\langle\mathbf{p} \mid \psi\rangle$ by the Fourier integral

$$
\begin{equation*}
\psi(\mathbf{x})=\int \frac{d^{3} p}{(2 \pi \hbar)^{3}} \tilde{\psi}(\mathbf{p}) e^{i \mathbf{p} \cdot \mathbf{x} / \hbar} \tag{4.47}
\end{equation*}
$$

which has the inverse

$$
\begin{equation*}
\tilde{\psi}(\mathbf{p})=\int d^{3} x \psi(\mathbf{x}) e^{-i \mathbf{p} \cdot \mathbf{x} / \hbar} \tag{4.48}
\end{equation*}
$$

in analogy with (4.33).
Operator matrix elements will follow from the generalizations of (4.26),

$$
\begin{equation*}
\langle\mathbf{x}| F(\widehat{\mathbf{x}}, \widehat{\mathbf{p}})|\psi\rangle=F\left(\mathbf{x},-i \hbar \partial_{\mathbf{x}}\right) \psi(\mathbf{x}) \tag{4.49}
\end{equation*}
$$

where $\partial_{\mathbf{x}}$ is the ordinary nabla operator with components $\partial_{\mathbf{x}}=\left(\partial_{x}, \partial_{y}, \partial_{z}\right)$. Similarly, in momentum space we now have

$$
\begin{equation*}
\langle\mathbf{p}| F(\widehat{\mathbf{x}}, \widehat{\mathbf{p}})|\psi\rangle=F\left(i \hbar \partial_{\mathbf{p}}, \mathbf{p}\right) \tilde{\psi}(\mathbf{p}) \tag{4.50}
\end{equation*}
$$

from (4.35).
In each direction we have the canonical commutars (4.27). Since operators acting in different directions, commute with each other, we can sum up all the three-dimensional canonical commutators in the single equation

$$
\begin{equation*}
\left[\widehat{x}_{i}, \widehat{p}_{j}\right]=i \hbar \delta_{i j} \tag{4.51}
\end{equation*}
$$

We will make much use of it in the following.

### 4.7 The wave equation

A particle moving in three dimensions will in general have some Hamiltonian operator $\widehat{H}=H(\widehat{\mathbf{x}}, \widehat{\mathbf{p}})$. It is described by a time-dependent state vector $|\Psi, t\rangle$ which is goverened by the Schrödinger equation

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t}|\Psi, t\rangle=H(\widehat{\mathbf{x}}, \widehat{\mathbf{p}})|\Psi, t\rangle \tag{4.52}
\end{equation*}
$$

For actual calculations of the time development we need a basis in Hilbert space where the state vector moves. Most used is the coordinate basis and the corresponding wavefunctions

$$
\begin{equation*}
\Psi(\mathbf{x}, t)=\langle\mathbf{x} \mid \Psi, t\rangle \tag{4.53}
\end{equation*}
$$

Now projecting the above Schrödinger equation onto the basis vector $|\mathbf{x}\rangle$, we get

$$
i \hbar \frac{\partial}{\partial t} \Psi(\mathbf{x}, t)=\langle\mathbf{x}| H(\widehat{\mathbf{x}}, \widehat{\mathbf{p}})|\Psi, t\rangle
$$

The matrix element on the right-hand side now follows from (4.49) resulting in the differential equation

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t} \Psi(\mathbf{x}, t)=H\left(\mathbf{x},-i \hbar \partial_{\mathbf{x}}\right) \Psi(\mathbf{x}, t) \tag{4.54}
\end{equation*}
$$

For a non-relativistic particle with mass $m$ moving in a potential $V(\mathbf{x})$ the Hamiltonian is

$$
\begin{equation*}
\widehat{H}=\frac{1}{2 m} \widehat{\mathbf{p}}^{2}+V(\widehat{\mathbf{x}}) \tag{4.55}
\end{equation*}
$$

resulting in the differential wave equation

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t} \Psi(\mathbf{x}, t)=\left[-\frac{\hbar^{2}}{2 m} \nabla^{2}+V(\mathbf{x})\right] \Psi(\mathbf{x}, t) \tag{4.56}
\end{equation*}
$$

This is obviously just one special example of the more general Schrödinger equation above.

There will also be a wavefunction $\tilde{\Psi}(\mathbf{p}, t)=\langle\mathbf{p} \mid \Psi, t\rangle$ in momentum space. Using now the result (4.50), we see that it will satisfy a corresponding differential equation. With the standard Hamiltonian (4.55) it will be

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t} \tilde{\Psi}(\mathbf{p}, t)=\left[\frac{1}{2 m} \mathbf{p}^{2}+V\left(i \hbar \partial_{\mathbf{p}}\right)\right] \tilde{\Psi}(\mathbf{p}, t) \tag{4.57}
\end{equation*}
$$

which is the Schrödinger wave equation in momentum space. For some special problems this is easier to solve than the coordinate wave equation (4.56).

These wavefunctions now determine the expectation values of different physical observables for this system. These will then be vary according to Ehrenfest's equation derived in Chapter 2. In particular, the expectation values $\langle\widehat{\mathbf{x}}\rangle$ and $\langle\widehat{\mathbf{p}}\rangle$ are now easy to calculate. In the first we need the commutator $\left[\hat{p}^{2}, \widehat{x}\right]=-2 i \hbar \widehat{p}$ from (4.28). Thus we get from (2.31)

$$
\begin{equation*}
\frac{d}{d t}\langle\widehat{\mathbf{x}}\rangle=\frac{1}{m} \widehat{\mathbf{p}} \tag{4.58}
\end{equation*}
$$

which is the quantum analogue of the classical relation between velocity and momentum. Similarly, from (4.31) we have $[V(\widehat{\mathbf{x}}), \widehat{\mathbf{p}}]=i \hbar V^{\prime}(\widehat{\mathbf{x}})$ and therefore

$$
\begin{equation*}
\frac{d}{d t}\langle\widehat{\mathbf{p}}\rangle=-V^{\prime}(\widehat{\mathbf{x}}) \tag{4.59}
\end{equation*}
$$

This is the quantum version of Newton's second law since $-V^{\prime}(\mathbf{x})$ is the classical force acting on the particle. We see here an aspect of what Bohr originally called the correspondence principle - the expectation values of quantum operators follow the classical equations of motion.

### 4.8 Algebraic solution of the harmonic oscillator

The Hamiltonian for a classical particle with mass $m$ moving in a one-dimensional harmonic oscillator potential with frequency $\omega$ is

$$
\begin{equation*}
H=\frac{p^{2}}{2 m}+\frac{1}{2} m \omega^{2} q^{2} \tag{4.60}
\end{equation*}
$$

Since it is quadratic in the coordinate $q$ and momentum $p$, it can be factorized as

$$
H=\omega\left(\sqrt{\frac{m \omega}{2}} q-i p \sqrt{\frac{1}{2 m \omega}}\right)\left(\sqrt{\frac{m \omega}{2}} q+i p \sqrt{\frac{1}{2 m \omega}}\right)
$$

This suggests that the in the corresponding quantum Hamiltonian

$$
\begin{equation*}
\widehat{H}=\frac{1}{2 m} \widehat{p}^{2}+\frac{1}{2} m \omega^{2} \widehat{q}^{2} \tag{4.61}
\end{equation*}
$$

a similar combination of the two operators $\widehat{q}$ and $\widehat{p}$ will be adventagious. And this is indeed very much the case. If we define the operator

$$
\begin{equation*}
\widehat{a}=\sqrt{\frac{m \omega}{2 \hbar}}\left(\widehat{q}+\frac{i}{m \omega} \widehat{p}\right) \tag{4.62}
\end{equation*}
$$

and the adjoint

$$
\begin{equation*}
\widehat{a}^{\dagger}=\sqrt{\frac{m \omega}{2 \hbar}}\left(\widehat{q}-\frac{i}{m \omega} \widehat{p}\right) \tag{4.63}
\end{equation*}
$$

we find by direct multiplication that

$$
\widehat{a}^{\dagger} \widehat{a}=\frac{1}{2 m \hbar \omega} \hat{p}^{2}+\frac{m \omega}{2 \hbar} \widehat{q}^{2}+\frac{i}{2 \hbar}[\widehat{q}, \widehat{p}]
$$

But the last term is non-zero and equal to the canonical commutator (4.27). So instead of the above classical factorization of the Hamiltonian, we now have

$$
\begin{equation*}
\widehat{a}^{\dagger} \widehat{a}=\frac{\widehat{H}}{\hbar \omega}-\frac{1}{2} \tag{4.64}
\end{equation*}
$$

or

$$
\begin{equation*}
\widehat{H}=\left(\widehat{a}^{\dagger} \widehat{a}+\frac{1}{2}\right) \hbar \omega \tag{4.65}
\end{equation*}
$$

Multiplying together the two operators in the opposite order, it similarly follows that

$$
\begin{equation*}
\widehat{a} \widehat{a}^{\dagger}=\frac{\widehat{H}}{\hbar \omega}+\frac{1}{2} \tag{4.66}
\end{equation*}
$$

Therefore, by subtracting these two products, one obtains the commutator

$$
\begin{equation*}
\left[\widehat{a}, \widehat{a}^{\dagger}\right]=1 \tag{4.67}
\end{equation*}
$$

which replaces the canonical commutator. The original operators $\widehat{q}$ and $\widehat{p}$ are now given by

$$
\begin{equation*}
\widehat{q}=\sqrt{\frac{\hbar}{2 m \omega}}\left(\widehat{a}^{\dagger}+\widehat{a}\right) \tag{4.68}
\end{equation*}
$$

and

$$
\begin{equation*}
\widehat{p}=i \sqrt{\frac{m \hbar \omega}{2}}\left(\widehat{a}^{\dagger}-\widehat{a}\right) \tag{4.69}
\end{equation*}
$$

Any calculation involving them can therefore be performed using the basic properties of these ladder operators.

We can now establish an energy basis in the Hilbert space of the oscillator by solving the energy eigenvalue problem

$$
\begin{equation*}
\widehat{H}\left|E_{n}\right\rangle=E_{n}\left|E_{n}\right\rangle \tag{4.70}
\end{equation*}
$$

in a purely algebraic way. It is equvalent to finding the eigenvalues of the so-called number operator $\widehat{n}=\widehat{a}^{\dagger} \widehat{a}$. In order to understand why this name arises, calculate the commutator $[\widehat{n}, \widehat{a}]$ or

$$
\left[\widehat{a}^{\dagger} \widehat{a}, \widehat{a}\right]=\widehat{a}^{\dagger}[\widehat{a}, \widehat{a}]+\left[\widehat{a}^{\dagger}, \widehat{a}\right] \widehat{a}=-\widehat{a}
$$

The state $\widehat{a}|n\rangle$ is now a new eigenstate of the Hamiltonian,

$$
\begin{aligned}
\widehat{H} \widehat{a}\left|E_{n}\right\rangle & =([\widehat{H}, \widehat{a}]+\widehat{a} \widehat{H})\left|E_{n}\right\rangle \\
& =\left(-\hbar \omega \widehat{a}+\widehat{a} E_{n}\right)\left|E_{n}\right\rangle=\left(E_{n}-\hbar \omega\right) \widehat{a}\left|E_{n}\right\rangle
\end{aligned}
$$

with energy $E_{n}-\hbar \omega$. We say that $\hat{a}$ acts as a lowering operator. Similarly, from $\left[\widehat{n}, \widehat{a}^{\dagger}\right]=\widehat{a}^{\dagger}$ it follows that the state $\widehat{a}^{\dagger}\left|E_{n}\right\rangle$ is also a new eigenstate with increased energy $E_{n}+\hbar \omega$. It is therefore said to act like a raising operator. Together they are ladder operators, allowing us to go up or down in energy in steps of $\hbar \omega$. This amount is called a quantum of energy.
By the same arguments the state $\widehat{a}^{2}\left|E_{n}\right\rangle$ will be an eigenstate of $\widehat{H}$ with eigenvalue $E_{n}-2 \hbar \omega$, i.e. reduced by two quanta. Continued application of the lowering operator could then eventually result in a negative energy. But the harmonic oscillator cannot have negative energy eigenstates. This follows from the fact that the norm of the state $\widehat{a}\left|E_{n}\right\rangle$ is positive. Therefore

$$
\left\langle E_{n}\right| \widehat{a}^{\dagger} \widehat{a}\left|E_{n}\right\rangle=\left\langle E_{n}\right| \frac{\widehat{H}}{\hbar \omega}-\frac{1}{2}\left|E_{n}\right\rangle=\frac{E_{n}}{\hbar \omega}-\frac{1}{2} \geq 0
$$

when the eigenstate $\left|E_{n}\right\rangle$ is normalized to one. Thus we must have $E_{n} \geq \frac{1}{2} \hbar \omega$. In order to avoid ending up with these forbidden, negative energies, it is necessary that in the chain of states generated by $\widehat{a}$, the state $\left|E_{0}\right\rangle$ corresponding to the lowest energy, is annihilated when the operator $\widehat{a}$ acts on it. It is said to be an empty state since it does not contain any quanta. For this reason we can just as well denote it by $|0\rangle$ - which should not be confused with the zero vector mentioned in the first chapters. In other words, we must have

$$
\begin{equation*}
\widehat{a}|0\rangle=0 \tag{4.71}
\end{equation*}
$$

From the Hamiltonian (4.65) this lowest state has a non-zero ground state energy

$$
\begin{equation*}
E_{0}=\frac{1}{2} \hbar \omega \tag{4.72}
\end{equation*}
$$

which is also often called the zero-point energy. It is a quantum effect, coming basically from the fact that the operators $\widehat{q}$ and $\widehat{p}$ don't commute with each other.

The first excited state is then $\left|E_{1}\right\rangle=\widehat{a}^{\dagger}|0\rangle$ with energy (3/2) $\hbar \omega$. In our slightly new notation, we denote it by $|1\rangle$. Using the raising operator on this ground state, we can then similarly construct the $n^{\prime}$ th excited state $|n\rangle$ with the energy

$$
\begin{equation*}
E_{n}=\left(n+\frac{1}{2}\right) \hbar \omega \tag{4.73}
\end{equation*}
$$

It contains $n$ quanta, each of energy $\hbar \omega$. Since now

$$
\begin{equation*}
\widehat{n}|n\rangle=n|n\rangle \tag{4.74}
\end{equation*}
$$

we say that the operator $\widehat{n}$ counts the number of quanta in the state,
The action of the ladder operators can be stated a bit more accurately. For instance, we must have $\widehat{a}|n\rangle=A_{n}|n-1\rangle$ where the factor $A_{n}$ is needed if the states on both sides of the equation are normalized to one. Calculating now the norm, we find $\left|A_{n}\right|^{2}=$ $\langle n| \widehat{a}^{\dagger} \widehat{a}|n\rangle=n$. Therefore

$$
\begin{equation*}
\widehat{a}|n\rangle=\sqrt{n}|n-1\rangle \tag{4.75}
\end{equation*}
$$

when we choose the phase factor to be one. In the same way from the norm of $\widehat{a}^{\dagger}|n\rangle$ follows

$$
\begin{equation*}
\widehat{a}^{\dagger}|n\rangle=\sqrt{n+1}|n+1\rangle \tag{4.76}
\end{equation*}
$$

These two results are vey useful in practical calculations. For instance, we can now find a closed expression for the $n^{\prime}$ th excited state built up directly from the ground state using

$$
\begin{align*}
|n\rangle & =\sqrt{\frac{1}{n}} \widehat{a}^{\dagger}|n-1\rangle=\sqrt{\frac{1}{n \cdot(n-1)}}\left(\widehat{a}^{\dagger}\right)^{2}|n-2\rangle  \tag{4.77}\\
& =\sqrt{\frac{1}{n!}}\left(\widehat{a}^{\dagger}\right)^{n}|0\rangle \tag{4.78}
\end{align*}
$$

We thus have algebraic results for all the energy eigenstates and the corresponding eigenvalues. Using this complete set of states as a basis in Hilbert space, we can calculating all matrix representations of operators involving the original $\widehat{q}$ and $\widehat{p}$ operators. Needless to stay, these will be infinite-dimensional matrices with indices given by the natural numbers from zero to infinity.
In the case of the three-dimensional oscillator with Hamiltonian

$$
\begin{equation*}
\widehat{H}=\frac{1}{2 m} \widehat{\mathbf{p}}^{2}+\frac{1}{2} m \omega^{2} \widehat{\mathbf{x}}^{2} \tag{4.79}
\end{equation*}
$$

we can find quantized solutions using the same algebraic approach. Introduce the lowering operators

$$
\begin{equation*}
\widehat{a}_{k}=\sqrt{\frac{m \omega}{2 \hbar}} \widehat{x}_{k}+i \sqrt{\frac{\hbar}{2 m \omega}} \widehat{p}_{k} \tag{4.80}
\end{equation*}
$$

and the corresponding adjoint raising operators $\widehat{a}_{k}^{\dagger}$, the Hamiltonian becomes

$$
\begin{equation*}
\widehat{H}=\left(\widehat{a}_{x}^{\dagger} \widehat{a}_{x}+\widehat{a}_{y}^{\dagger} \widehat{a}_{y}+\widehat{a}_{z}^{\dagger} \widehat{a}_{z}+\frac{3}{2}\right) \hbar \omega \tag{4.81}
\end{equation*}
$$

These new ladder operators commute with each other except for

$$
\begin{equation*}
\left[\widehat{a}_{i}, \widehat{a}_{j}^{\dagger}\right]=\delta_{i j} \tag{4.82}
\end{equation*}
$$

The operators $\widehat{N}_{k}=\widehat{a}_{k}^{\dagger} \widehat{a}_{k}$ thus give the numbers of quanta excited in the three directions, i.e. $n_{x}, n_{y}$ and $n_{z}$. In this way we obtain for the eigenvalues of the Hamiltonian (4.81)

$$
\begin{equation*}
E_{n_{x} n_{y} n_{z}}=\left(n_{x}+n_{y}+n_{z}+\frac{3}{2}\right) \hbar \omega \tag{4.83}
\end{equation*}
$$

From (4.78) follows the corresponding eigenstates as

$$
\begin{equation*}
\left|n_{x} n_{y} n_{z}\right\rangle=\left|n_{x}\right\rangle\left|n_{y}\right\rangle\left|n_{z}\right\rangle=\sqrt{\frac{1}{n_{x}!n_{y}!n_{z}!}}\left(\widehat{a}_{x}^{\dagger}\right)^{n_{x}}\left(\widehat{a}_{y}^{\dagger}\right)^{n_{y}}\left(\widehat{a}_{z}^{\dagger}\right)^{n_{z}}|0\rangle \tag{4.84}
\end{equation*}
$$

Notice that the excited states are now degenerate. The first excited energy level with energy $(5 / 2) \hbar \omega$ is triple degenerate consisting of the three states $|100\rangle,|010\rangle$ and $|001\rangle$. Similarly we find that the second excited level is six-fold degenerate and so on.
So far we have worked in the Schrödinger picture where all operators are independent of time. However, many calculations and in particular connections with classical mechanics are simpler in the Heisenberg picture. Since all operators for the oscillator can be expressed in terms of ladder operators, we need to find their time dependence in the Heisenberg picture. It will follow from the basic Heisenberg equation (2.48) derived in Chapter 2,

$$
\frac{d}{d t} \widehat{a}(t)=\frac{i}{\hbar}[\widehat{H}, \widehat{a}]
$$

Since here $\widehat{H}=\hbar \omega(\widehat{N}+1 / 2)$, it simply gives

$$
\frac{d}{d t} \widehat{a}(t)=-i \omega \widehat{a}(t)
$$

or

$$
\begin{equation*}
\widehat{a}(t)=\widehat{a} e^{-i \omega t} \tag{4.85}
\end{equation*}
$$

when integrated. The creation operator $\widehat{a}^{\dagger}$ has similarly the complex conjugate time dependence. We could have found these results just as easily using the basic transformation (2.47) to the Heisenberg picture combined with the Lie formula (2.53) which was found in Chapter 2. Either way, for the position operator (4.68) we now have

$$
\begin{equation*}
\widehat{q}(t)=\sqrt{\frac{\hbar}{2 m \omega}}\left(\widehat{a} e^{-i \omega t}+\widehat{a}^{\dagger} e^{i \omega t}\right) \tag{4.86}
\end{equation*}
$$

Surprisingly enough, this expression will appear again when we quantize classical fields as for electromagnetism in in Chapter 7.

### 4.9 Hermite polynomials

The energy eigenstates (4.78) will in a coordinate representation be given by the eigenfunctions $\psi_{n}(q)=\langle q \mid n\rangle$ satisfying the differential wave equation

$$
\begin{equation*}
\left(-\frac{\hbar^{2}}{2 m} \frac{d^{2}}{d q^{2}}+\frac{1}{2} m \omega^{2} q^{2}\right) \psi_{n}(q)=E_{n} \psi_{n}(q) \tag{4.87}
\end{equation*}
$$

This is just the projection of (4.70) onto the coordinate basis. Using the general result (4.26), the wavefunctions will now given by

$$
\begin{aligned}
\psi_{n}(q) & =\sqrt{\frac{1}{n!}}\langle q|\left(\widehat{a}^{\dagger}\right)^{n}|n\rangle \\
& =\sqrt{\frac{1}{n!}}\left(\sqrt{\frac{m \omega}{2 \hbar}} q-\sqrt{\frac{\hbar}{2 m \omega}} \frac{d}{d q}\right)^{n}\langle q \mid 0\rangle
\end{aligned}
$$

On the right-hand side we find $\psi_{0}(q)=\langle q \mid 0\rangle$ which is the wave function of the ground state. Introducing $\alpha=\sqrt{m \omega / \hbar}$ and the dimensional coordinate $x=\alpha q$, we have

$$
\begin{equation*}
\psi_{n}(q)=\sqrt{\frac{1}{2^{n} n!}}\left(x-\frac{d}{d x}\right)^{n} \psi_{0}(q) \tag{4.88}
\end{equation*}
$$

All the wavefunctions for excited states can in this way we calculated from knowledge of the ground state wavefunction. This latter one follows from the definition (4.71) which gives $\langle q| \widehat{a}|0\rangle=0$ or

$$
\left(\sqrt{\frac{m \omega}{2 \hbar}} q+\sqrt{\frac{\hbar}{2 m \omega}} \frac{d}{d q}\right) \psi_{0}(q)=0
$$

This is a simple differential equation with the solution

$$
\psi_{0}(q)=C e^{-m \omega q^{2} / 2 \hbar}
$$

Here $C$ is an integration constant which is fixed by the normalization

$$
\langle 0 \mid 0\rangle=\int_{-\infty}^{\infty} d q\left|\psi_{0}(q)\right|^{2}=1
$$

giving $|C|^{2}=\sqrt{m \omega / \pi \hbar}$. The ground state function is therefore

$$
\begin{equation*}
\psi_{0}(q)=\left(\frac{m \omega}{\pi \hbar}\right)^{1 / 4} e^{-m \omega^{2} q^{2} / 2 \hbar} \tag{4.89}
\end{equation*}
$$

when we again set an arbitrary phase factor equal to unity.
The result (4.88) for the excited wavefunctions now can be written as

$$
\begin{equation*}
\psi_{n}(q)=\left(\frac{m \omega}{\pi \hbar}\right)^{1 / 4} \sqrt{\frac{1}{2^{n} n!}} H_{n}(x) e^{-x^{2} / 2} \tag{4.90}
\end{equation*}
$$

where the functions

$$
\begin{equation*}
H_{n}(x)=e^{x^{2} / 2}\left(x-\frac{d}{d x}\right)^{n} e^{-x^{2} / 2}=e^{x^{2}}\left(-\frac{d}{d x}\right)^{n} e^{-x^{2}} \tag{4.91}
\end{equation*}
$$

are the Hermite polynomials. From this formula of Rodriquez follow the lowest ones as

$$
\begin{align*}
H_{0}(x) & =1 \\
H_{1}(x) & =2 x \\
H_{2}(x) & =4 x^{2}-2  \tag{4.92}\\
H_{3}(x) & =8 x^{3}-12 x
\end{align*}
$$

We see that these are even functions when $n$ is even and odd functions when $n$ is odd. This just gives the parity quantum number under the reflection $x \rightarrow-x$ corresponding to the act of the parity operator discussed previously.

### 4.10 Coherent states

When the oscillator is in an energy eigenstate, the motion is stationary and there is very little to be seen of the classical oscillations with fixed frequency $\omega$. But we will here show that a certain, infinite superposition of such states behave in many ways as a clasical state. This is the coherent state defined as an eigenstate of the annihilation operator, i.e.

$$
\begin{equation*}
\widehat{a}|\alpha\rangle=\alpha|\alpha\rangle \tag{4.93}
\end{equation*}
$$

Since $\widehat{a}$ is not an Hermitean operator, the eigenvalue $\alpha$ will in general be complex. The eigenstate $|\alpha\rangle$ must be of the form

$$
|\alpha\rangle=\sum_{n=0}^{\infty} c_{n}|n\rangle
$$

where the components $c_{n}$ can be determined from the above eigenvalue equation. One finds, using (4.75),

$$
\widehat{a}|\alpha\rangle=\sum_{n=0}^{\infty} c_{n} \widehat{a}|n\rangle=\sum_{n=1}^{\infty} c_{n} \sqrt{n}|n-1\rangle
$$

After the shift $n-1 \rightarrow n$ of summation variable, this becomes

$$
\widehat{a}|\alpha\rangle=\sum_{n=0}^{\infty} c_{n+1} \sqrt{n+1}|n\rangle
$$

Comparing with the original expansion, we thus have obtained the recursion relation $c_{n+1}=\alpha c_{n} / \sqrt{n+1}$. Iterating, it gives

$$
c_{n+1}=\frac{\alpha^{n+1}}{\sqrt{n+1}!} c_{0}
$$

and therefore for the coherent state

$$
\begin{equation*}
|\alpha\rangle=c_{0} \sum_{n=0}^{\infty} \frac{\alpha^{n}}{\sqrt{n!}}|n\rangle=c_{0} \sum_{n=0}^{\infty} \frac{\alpha^{n}}{n!}\left(\widehat{a}^{\dagger}\right)^{n}|0\rangle=c_{0} e^{\alpha \widehat{a}^{\dagger}}|0\rangle \tag{4.94}
\end{equation*}
$$

It contains non-zero amplitudes for being in any of the excited states of the oscillator. The component $c_{0}$ is the probability amplitude for finding the the oscillator in the ground state when this coherent state is normalized to $\langle\alpha \mid \alpha\rangle=1$. Now we have

$$
\langle\alpha \mid \alpha\rangle=\left|c_{0}\right|^{2}\langle 0| e^{\alpha^{*} \widehat{a}} e^{\alpha \widehat{a} \dagger}|0\rangle
$$

We can here make use of the Weyl formula at the end of Chapter 2. It can be written as

$$
\begin{equation*}
e^{\widehat{A}} e^{\widehat{B}}=e^{\widehat{B}} e^{\widehat{A}} e^{[\widehat{A}, \widehat{B}]} \tag{4.95}
\end{equation*}
$$

when the operators $\widehat{A}$ and $\widehat{B}$ both commute with their commutator. This applies to the above expectation value which then gives $\left|c_{0}\right|^{2} \exp \left(|\alpha|^{2}\right)=1$ or

$$
\begin{equation*}
c_{0}=e^{-|\alpha|^{2} / 2} \tag{4.96}
\end{equation*}
$$

when we set the phase factor equal to one. From (4.94) we can then read off the probability

$$
\begin{equation*}
P_{n}=e^{-|\alpha|^{2}} \frac{|\alpha|^{2 n}}{n!} \tag{4.97}
\end{equation*}
$$

to find $n$ quanta in the coherent state. This is just the Poisson distribution.
Applying the time evolution operator to the coherent state, it changes into the new state

$$
\begin{equation*}
|\alpha, t\rangle=e^{-i \widehat{H} t / \hbar}|\alpha\rangle=c_{0} \sum_{n=0}^{\infty} \frac{\alpha^{n}}{\sqrt{n!}} e^{-i \widehat{H} t / \hbar}|n\rangle \tag{4.98}
\end{equation*}
$$

Since $|n\rangle$ is an eigenstate of the Hamiltonian $\widehat{H}$ with eigenvalue $E_{n}=(n+1 / 2) \hbar \omega$, we have

$$
\begin{aligned}
|\alpha, t\rangle & =c_{0} \sum_{n=0}^{\infty} \frac{\alpha^{n}}{\sqrt{n!}} e^{-i \omega t(n+1 / 2)}|n\rangle=c_{0} e^{-i \omega t / 2} \sum_{n=0}^{\infty} \frac{\left(\alpha e^{-i \omega t}\right)^{n}}{\sqrt{n!}}|n\rangle \\
& =e^{-i \omega t / 2}\left|\alpha e^{-i \omega t}\right\rangle
\end{aligned}
$$

It therefore remains an eigenstate

$$
\begin{equation*}
\widehat{a}|\alpha, t\rangle=\alpha e^{-i \omega t}|\alpha, t\rangle \tag{4.99}
\end{equation*}
$$

of the annihilation operator at all times, but the corresponding eigenvalue oscillates with the fundamental frequency.
It now instructive to calculate the expectation value $\langle\widehat{q}\rangle$ of the position of the particle in the oscillator potential. Using the relation (4.68), it is

$$
\langle\widehat{q}\rangle=\sqrt{\frac{\hbar}{2 m \omega}}\langle\alpha, t| \widehat{a}+\widehat{a}^{\dagger}|\alpha, t\rangle=\sqrt{\frac{\hbar}{2 m \omega}}\left(\alpha e^{-i \omega t}+\alpha^{*} e^{i \omega t}\right)
$$

from (4.99). Writing now for the complex parameter $\alpha=|\alpha| e^{i \theta}$, we finally have

$$
\begin{equation*}
\langle\widehat{q}\rangle=\sqrt{\frac{2 \hbar}{m \omega}}|\alpha| \cos (\omega t-\theta) \tag{4.100}
\end{equation*}
$$

This is just the classical, oscillatory motion with an amplitude proportional to the modulus $|\alpha|$ of the coherent parameter.

## Chapter 5

## Rotations and spherical symmetry

Systems in two and three dimensions can not only be translated, but also rotated. In the same way as the momentum operator arises as the generator of translations, we will now see that the angular momentum or spin operator is the generator of rotations. When the system has spherical symmetry, it is invariant under these transformations and the angular momentum of the system is conserved. Quantization gives only discrete values for this new spin quantum number. Of special interest will be systems with the lowest spin value $1 / 2$ described by the Pauli matrices. Before considering ordinary rotations in three dimensions, we will first consider the simpler situation with rotations in the plane.

### 5.1 Cyclic molecules

Before we consider the ordinary continuous rotations in space, it is instructive to take a closer look at discrete rotations in the plane. This has also direct applications to cyclic molecules which are important in chemistry. Such a system is most easily constructed from the one-dimensional lattice considered in Chapter 3 when it contains a finite number $N$ of sites or atoms and the last atom is connected to the first so that it forms a ring. An example of this is shown in Fig.5.1 which applies to the organic molecule benzene. We will again consider a quantum particle on this lattice, assuming that it can only be localized to one of the $N$ different sites. This is what the chemists call the Hückel approximation for such molecules.
There will now be $N$ orthonormal basis states $|n\rangle$ where the cyclic property of the chain requires $|n+N\rangle=|n\rangle$. Instead of the distance $a$ between nearest-neighbour lattice sites which was called the lattice constant in Chapter 3, it is now more convenient to describe this separation by the angle $\alpha$. From the figure we see that it has the magnitude $\alpha=2 \pi / N$. For the same reason, we introduce the unitary rotation operator $\widehat{R}(\alpha)$ defined by

$$
\begin{equation*}
\widehat{R}(\alpha)|n\rangle=|n+1\rangle \tag{5.1}
\end{equation*}
$$

which replaces the fundamental translation operator $\widehat{T}(a)$ defined in (3.28). The product of $p$ such elementary rotations after each other will be equivalent to a single rotation by an angle $\phi=p \alpha$ so that $\widehat{R}^{p}(\alpha)|n\rangle=\widehat{R}(\phi)|n\rangle=|n+p\rangle$.


Figure 5.1: Cyclic molecule with $N=6$ sites or atoms. It has the same structure as benzene.

The molecule is invariant under these rotations. In the quantum description there will be corresponding rotationally invariant states. These we now define as

$$
\begin{equation*}
\widehat{R}(\alpha)|\psi\rangle=e^{-i m \alpha}|\psi\rangle \tag{5.2}
\end{equation*}
$$

in analogy with (3.32). Here $m$ is a new quantum number which so far can take on any real value. Repeating these elementary rotations $p$ times, we find

$$
\begin{equation*}
\widehat{R}(\phi)|\psi\rangle=e^{-i m \phi}|\psi\rangle \tag{5.3}
\end{equation*}
$$

for a rotation angle $\phi=p \alpha$.
Expanding the state in the site basis, we have

$$
\begin{equation*}
|\psi\rangle=\sum_{n=1}^{N} \psi_{n}|n\rangle \tag{5.4}
\end{equation*}
$$

where again the coefficient $\psi_{n}$ is the probability amplitude to find the particle on site $n$. From the definition in (5.2) now follows that this amplitude is

$$
\begin{equation*}
\psi_{n}=e^{i m n \alpha} \tag{5.5}
\end{equation*}
$$

with an appropriate normalization. But since the atom at position $n+N$ is the atom at $n$, we now have the cyclic requirement $\psi_{n+N}=\psi_{n}$ which gives the condition $\exp (\operatorname{imN\alpha })=$ 1. With the above value for the angle $\alpha$, we thus get $\exp (2 \pi i m)=1$ or $m=1,2,3, \ldots, N$. Other positive or negative values will just reproduce the same states.

In order to calculate the energy of the particle in these rotation eigenstates, we need the Hamiltonian. Assuming again as in Chapter 3 only quantum jumps between nearestneighbour sites, it is

$$
\begin{equation*}
\widehat{H}=E_{0} \widehat{I}-A\left(\widehat{R}+\widehat{R}^{\dagger}\right) \tag{5.6}
\end{equation*}
$$

where $E_{0}$ is the energy of the particle localized to one particular site and $A$ gives the amplitude to jump. The energy eigenvalues now follow from $\widehat{H}|\psi\rangle=E|\psi\rangle$ which gives

$$
\begin{equation*}
E=E_{0}-2 A \cos m \alpha=E_{0}-2 A \cos \frac{2 \pi m}{N} \tag{5.7}
\end{equation*}
$$

when the particle is in the state with rotational quantum number $m$ as in the corresponding result (3.36) for translations. Assuming $A>0$, the state with lowest energy has $m=N$ which is equivalent to $m=0$. For benzene with $N=6$ this becomes $E=E_{0}-2 A \cos (\pi m / 3)$. As shown in Fig.5.2 we can geometrically presented these eigenvalues as the projections onto the vertical energy axis of the corner sites of an hexagon with radius $2 A$ and centered at $E_{0}$. This is called the dual or reciprocal lattice


Figure 5.2: Cyclic molecule with $N$ sites has $N$ eigenstates on reciprocal lattice.
of the original lattice of the molecule. It is seen to have the same cyclic symmetry. The ground state has energy $E_{0}-2 A$, there are two excited states with degenerate energies $E_{0}-A$, two with energies $E_{0}+A$ and one with the highest energy $E_{0}+2 A$.

### 5.2 Particle on a circle

In the continuous limit where $N \rightarrow \infty$ and thus $\alpha \rightarrow 0$, our cyclic molecule becomes a circle. But the rotation operator $\widehat{R}(\phi)$ will still act as in (5.3) on a rotationally invariant state with the only difference that the quantum number $m$ now can take on infinitely many different values $m=0, \pm 1, \pm 2, \ldots, \pm \infty$. Introducing the new notation $|\psi\rangle \rightarrow|m\rangle$ for these eigenstates, we therefore have

$$
\begin{equation*}
\widehat{R}(\phi)|m\rangle=e^{-i m \phi}|m\rangle \tag{5.8}
\end{equation*}
$$

An infinitesemal rotation by an angle $\delta \phi$ will be slightly different from the unit operator and we write it as

$$
\begin{equation*}
\widehat{R}(\delta \phi)=\widehat{I}-\frac{i}{\hbar} \delta \phi \widehat{p}_{\phi} \tag{5.9}
\end{equation*}
$$

where now $\widehat{p}_{\phi}$ is the generator of rotations. It is a Hermitean operator called angular momentum and corresponds to the linear momentum operator $\widehat{p}_{x}$ in Chapter 3. By iterations we then get the full operator

$$
\begin{equation*}
\widehat{R}(\phi)=e^{-i \phi \widehat{p}_{\phi} / \hbar} \tag{5.10}
\end{equation*}
$$

for a finite rotation by an angle $\phi$. Comparing with (5.8) we see that the eigenstate $|m\rangle$ is also an eigenstate of this new generator,

$$
\begin{equation*}
\widehat{p}_{\phi}|m\rangle=\hbar m|m\rangle \tag{5.11}
\end{equation*}
$$

i.e. with the eigenvalue $p_{\phi}=\hbar m$. Thus the angular momentum for a particle on a circle is quantized with these integral values. The corresponding wave function of the particle is now

$$
\begin{equation*}
\psi_{m}(\phi)=\langle\phi \mid m\rangle=e^{i m \phi} \tag{5.12}
\end{equation*}
$$

which is seen to satisfy periodicity $\psi_{m}(\phi+2 \pi)=\psi_{m}(\phi)$ as it should. In this coordinate representation, or rather angle representation, the angular momentum operator becomes the differential operator

$$
\begin{equation*}
\widehat{p}_{\phi} \rightarrow-i \hbar \frac{\partial}{\partial \phi} \tag{5.13}
\end{equation*}
$$

as we showed for linear momentum.
If the particle with mass $\mu$ moves on a circle with radius $a$ and angular momentum $p_{\phi}$, it would have the classical energy $E=p_{\phi}^{2} / 2 \mu a^{2}$. The quantum motion is therefore determined by the Hamiltonian operator

$$
\begin{equation*}
\widehat{H}=\frac{\widehat{p}_{\phi}^{2}}{2 \mu a^{2}} \tag{5.14}
\end{equation*}
$$

where $\widehat{p}_{\phi}$ is the above angular momentum operator. From the eigenfunctions (5.12) we find the corresponding discrete energy eigenvalues $E_{m}=\left(\hbar^{2} / 2 \mu a^{2}\right) m^{2}$. The ground state has the quantum number $m=0$ while every excited state is double degenerate.

We can make use of this result for particle having a wave function which is required to be periodic in a direction $x$ with period $L$. This is equivalent to having the particle on a circle with radius $R=L / 2 \pi$. Now since $\phi=x / R=2 \pi x / L$, we can write the wave function as

$$
\begin{equation*}
\psi_{k}(x)=e^{2 \pi i m x / L} \equiv e^{i k x} \tag{5.15}
\end{equation*}
$$

It is thus a plane wave with wave number quantized as $k=2 \pi m / L$. By construction, it satisfies $\psi_{k}(x+L)=\psi_{k}(x)$.

### 5.3 Axial symmetry

The operator $\widehat{p}_{\phi}$ generates rotations by an angle $\phi$ around an axis normal to the plane of motion for the particle. When it can move on a circle, the system is rotational invariant around this axis and is said to have axial symmetry or invariance. In a more general case the particle can move in the plane not confined to a circle, but by a potential $V(r)$ which only depends on the radial length $\left.r=x^{2}+y^{2}\right)^{1 / 2}$ and not the polar angle $\phi$. Obviously, the system still has axial symmetry.

If the particle again has mass $\mu$, the Hamiltonian will now be

$$
\begin{equation*}
H=-\frac{\hbar^{2}}{2 \mu} \nabla^{2}+V(r) \tag{5.16}
\end{equation*}
$$

in the coordinate representation. In polar coordinates $x=r \cos \phi$ and $y=r \sin \phi$ we have

$$
\begin{equation*}
\nabla^{2}=\frac{\partial^{2}}{\partial r^{2}}+\frac{1}{r} \frac{\partial}{\partial r}+\frac{1}{r^{2}} \frac{\partial^{2}}{\partial \phi^{2}} \tag{5.17}
\end{equation*}
$$

and the energy eigenvalues $E$ will be given by the stationary Schrödinger equation

$$
\begin{equation*}
\left[-\frac{\hbar^{2}}{2 \mu}\left(\frac{\partial^{2}}{\partial r^{2}}+\frac{1}{r} \frac{\partial}{\partial r}+\frac{1}{r^{2}} \frac{\partial^{2}}{\partial \phi^{2}}\right)+V(r)\right] \psi(r, \phi)=E \psi(r, \phi) \tag{5.18}
\end{equation*}
$$

From the structure of this differential equation for the eigenfunctions $\psi(r, \phi)$, we see that it allows us to separate the dependence on the variables $r$ and $\phi$ by writing $\psi(r, \phi)=$ $R(r) Y(\phi)$. It can then be rewritten as

$$
\frac{r^{2}}{R}\left[\frac{d^{2} R}{d r^{2}}+\frac{1}{r} \frac{d R}{d r}+\frac{2 \mu}{\hbar^{2}}(E-V) R\right]=-\frac{1}{Y} \frac{d^{2} Y}{d \phi^{2}}
$$

For this to be satsified for all values of the coordinates, each side of the equation must equal a constant which we call $-m^{2}$. From the right-hand side we then get the simple differential equation

$$
\begin{equation*}
\frac{d^{2} Y}{d \phi^{2}}=-m^{2} Y \tag{5.19}
\end{equation*}
$$

with solutions of the form $Y_{m}(\phi)=e^{i m \phi}$ when we ignore any normalization. For this to be invariant under the full rotation $\phi \rightarrow \phi+2 \pi$, we see that $m$ must be an integer, $m=0, \pm 1, \pm 2, \ldots, \infty$. It is therefore just the quantized values of the planar angular momentum. From our previous knowledge, we could have assumed this form of the angular part of the wave function right from the beginning.

The remaining, radial equation now becomes

$$
\begin{equation*}
-\frac{\hbar^{2}}{2 \mu}\left[\frac{d^{2} R}{d r^{2}}+\frac{1}{r} \frac{d R}{d r}-\frac{m^{2}}{r^{2}}\right] R=(E-V) R \tag{5.20}
\end{equation*}
$$

and can in general for an arbitrary potential $V(r)$ not be solved analytically. But in the case of a free particle, i.e. when $V=0$ it simplifies to a well-known equation. Introducing $k^{2}=2 \mu E / \hbar^{2}$, it can then be written as

$$
\begin{equation*}
x^{2} R^{\prime \prime}+x R^{\prime}+\left(x^{2}-m^{2}\right) R=0 \tag{5.21}
\end{equation*}
$$

where the dash represents derivation with respect to the dimensionless radial variable $x=k r$. This is Bessel's differential equation. Requiring the corresponding solutions to be regular as $x \rightarrow 0$, these are given by the Bessel functions $J_{m}(x)$. The full wave functions for a free particle are then of the form

$$
\begin{equation*}
\psi_{m}(r, \phi)=N_{m} e^{i m \phi} J_{m}(k r) \tag{5.22}
\end{equation*}
$$

where $N_{m}$ is a normalization constant. Except for $J_{0}(x)$ which is non-zero at $x=0$, the other Bessel function $J_{m}(x)$ are all zero there and have like $J_{0}(x)$ an infinite number of other zeros at the nodes $x_{m n}$. For $J_{0}(x)$ these are at $x_{01}=2.405, x_{02}=5.520, x_{03}=$ $8.654, \ldots$, for $J_{1}(x)$ at $x_{11}=3.832, x_{12}=7.016, \ldots$ and so on.
We can now make use of these simple results for the more physical situation of a particle confined within an infinite and circular well in the plane described by the potential

$$
V(r)= \begin{cases}0, & r<a  \tag{5.23}\\ \infty, & r>a\end{cases}
$$

where $a$ is the radius of the well. The radial wave function $R(r)$ must therefore be zero for $r=a$ resulting in the quantization condition $k a=x_{m n}$ for a solution with angular momentum $m$. Since the wave number $k$ is given by the energy, we thus find that the energy eigenvalues are

$$
\begin{equation*}
E_{m n}=\frac{\hbar^{2}}{2 \mu a^{2}} x_{m n}^{2} \tag{5.24}
\end{equation*}
$$

For each value of the angular quantum number $m$ there is therefore an infinite number of excited states labelled by the quantum number $n$. This also gives the number of zeros in the radial wave function. The higher the number of such roots is, the higher the energy of the corresponding state is.

### 5.4 Planar harmonic oscillator

We saw in Chapter 4 that the harmonic oscillator can be quantized algebraically in any dimension using Cartesian coordinates. In two dimensions where the harmonic potential is

$$
\begin{equation*}
V(x, y)=\frac{1}{2} \mu \omega^{2}\left(x^{2}+y^{2}\right)=\frac{1}{2} \mu \omega^{2} r^{2} \tag{5.25}
\end{equation*}
$$

we would therefore expect that the radial equation (5.20) also allows exact solutions. It now becomes

$$
\frac{d^{2} R}{d r^{2}}+\frac{1}{r} \frac{d R}{d r}-\frac{m^{2}}{r^{2}} R+\left(k^{2}-\lambda^{2} r^{2}\right) R=0
$$

where we have introduced $\lambda=\mu \omega / \hbar$. Again we want to find solutions which are regular when $r \rightarrow 0$. In this region the angular momentum term dominates and we have approximately

$$
\frac{d^{2} R}{d r^{2}}+\frac{1}{r} \frac{d R}{d r}-\frac{m^{2}}{r^{2}} R=0
$$

with the simple solution $R=r^{|m|}$. Similarly, for large distances $r \rightarrow \infty$ the potential term dominates and the equation becomes

$$
\frac{d^{2} R}{d r^{2}}+\frac{1}{r} \frac{d R}{d r}-\lambda^{2} r^{2} R=0
$$

Again there is a simple solution which is bounded in this region, namely $R=e^{-\lambda r^{2} / 2}$. We can now separate out these two asymptotic solutions from the general, radial function and write

$$
R(r)=r^{|m|} e^{-\lambda r^{2} / 2} u(r)
$$

where the unknown function $u(r)$ must be well-behaved both when $r \rightarrow 0$ and $r \rightarrow \infty$. Introducing now the dimensionless radial variable $t=\lambda r^{2}$, we can rewrite the above differential equation as

$$
\begin{equation*}
t u^{\prime \prime}+(1+|m|-t) u^{\prime}-\frac{1}{2}\left(1+|m|-\frac{k^{2}}{4 \lambda}\right) u=0 \tag{5.26}
\end{equation*}
$$

after a calculating a few derivatives in this new variable. This is Kummer's differential equation of the general form

$$
\begin{equation*}
x y^{\prime \prime}+(1+m-x) y^{\prime}-a y=0 \tag{5.27}
\end{equation*}
$$

For the function $y(x)$ not to blow up when $x \rightarrow \infty$, the coefficient $a=-n$ with $n=0,1,2, \ldots, \infty$. It then goes over into the Laguerre differential equation $x y^{\prime \prime}+(1+$ $m-x) y^{\prime}+n y=0$ with the Rodrigues solutions

$$
\begin{equation*}
y(x)=L_{n}^{m}(x)=\frac{1}{n!} e^{x} x^{-m}\left(\frac{d}{d x}\right)^{n}\left(x^{m+n} e^{-x}\right) \tag{5.28}
\end{equation*}
$$

called associated Laguerre polynomials. The ordinary Laguerre polynomials are $L_{n}(x)=$ $L_{n}^{0}(x)$ or

$$
\begin{equation*}
L_{n}(x)=\frac{1}{n!} e^{x}\left(\frac{d}{d x}\right)^{n}\left(x^{n} e^{-x}\right) \tag{5.29}
\end{equation*}
$$

so that they are polynomials of order $n$. One can demonstrate that

$$
\begin{equation*}
L_{n}^{m}(x)=\left(-\frac{d}{d x}\right)^{m} L_{m+n}(x) \tag{5.30}
\end{equation*}
$$

which shows that these polynomials are also of order $n$. The lowest Laguerre polynomials are

$$
\begin{aligned}
& L_{0}=1 \\
& L_{1}=1-x \\
& L_{2}=1-2 x+\frac{1}{2} x^{2} \\
& L_{3}=1-3 x+\frac{3}{2} x^{2}-\frac{1}{6} x^{3} \\
& L_{4}=1-4 x+3 x^{2}-\frac{2}{3} x^{3}+\frac{1}{24} x^{4}
\end{aligned}
$$

It then follows that $L_{0}^{m}(x)=1, L_{1}^{1}=2-x, L_{1}^{2}=3-x$ while $L_{2}^{2}=6-4 x+x^{2} / 2$ and so on.

Returning now to the two-dimensional oscillator described by the radial differential equation (5.26), it has regular solutions in terms of associated Laguerre polynomials
$L_{n}^{|m|}(t)$ when $1+|m|-k^{2} / 4 \lambda=-2 n$ where the radial quantum number $n=0,1,2, \ldots, \infty$ for every value of the angular momentum quantum number $m$. Expressing the parameter $k^{2}$ by the energy $E$, we then have the quantized energy eigenvalues

$$
\begin{equation*}
E_{m n}=(2 n+|m|+1) \hbar \omega \tag{5.31}
\end{equation*}
$$

while the corresponding unnormalized eigenstates are

$$
\begin{equation*}
\psi_{m n}(r, \phi)=e^{i m \phi} r^{|m|} e^{-\mu \omega r^{2} / 2 \hbar} L_{n}^{|m|}\left(\mu \omega r^{2} / \hbar\right) \tag{5.32}
\end{equation*}
$$

This should be compared with the results in Cartesian coordinates. The eigenfunctions were then products of Hermite polynomials for each direction while the eigenvalues became $E_{n_{x} n_{y}}=\hbar \omega\left(n_{x}+n_{y}+1\right)$. We can easily convince ourselves that this agrees with what we have just found in polar coordinates even when we check the degeneracy of each energy level.

We can link these two derivations of the spectrum for the planar oscillator by introducing the new operators

$$
\begin{equation*}
\widehat{a}_{ \pm}=\sqrt{\frac{1}{2}}\left(\widehat{a}_{x} \mp i \widehat{a}_{y}\right) \tag{5.33}
\end{equation*}
$$

having the non-zero commutators $\left[\widehat{a}_{ \pm}, \widehat{a}_{ \pm}^{\dagger}\right]=1$. Since the Cartesian ladder operators $\left(\widehat{a}_{x}, \widehat{a}_{y}\right)$ and their adjoints are lowering and raising operators for the energy, so are also these circular operators $\widehat{a}_{ \pm}$and their adjoints. In fact, we find that the Hamiltonian can be written as

$$
\begin{equation*}
\widehat{H}=\left(\widehat{a}_{+}^{\dagger} \widehat{a}_{+}+\widehat{a}_{-}^{\dagger} \widehat{a}_{-}+1\right) \hbar \omega \tag{5.34}
\end{equation*}
$$

so that $\left[\widehat{H}, \widehat{a}_{ \pm}\right]=-\widehat{a}_{ \pm}$and $\left[\widehat{H}, \widehat{a}_{ \pm}^{\dagger}\right]=\widehat{a}_{ \pm}^{\dagger}$. But they are also seen to change the angular momentum

$$
\begin{equation*}
\widehat{L}_{z}=\widehat{x} \widehat{p}_{y}-\widehat{y} \widehat{p}_{x}=\hbar\left(\widehat{a}_{+}^{\dagger} \widehat{a}_{+}-\widehat{a}_{-}^{\dagger} \widehat{a}_{-}\right) \tag{5.35}
\end{equation*}
$$

of the particle. So while both these circular raising operators increase the energy by $\hbar \omega$, it follows from the same reasoning that $\widehat{a}_{+}^{\dagger}$ increases the angular momentum by $\hbar$ while $\widehat{a}_{-}^{\dagger}$ decreases it by the same amount. Since we said the number operators $\widehat{a}_{x}^{\dagger} \widehat{a}_{x}$ and $\widehat{a}_{y}^{\dagger} \widehat{a}_{y}$ count the number of quanta in the $x$ and $y$ directions, we can say that $\widehat{a}_{+}^{\dagger} \widehat{a}_{+}$and $\widehat{a}_{-}^{\dagger} \widehat{a}_{-}$ count the number of 'up' quanta with spin along the $z$-axis and 'down' quanta with spin opposite to the $z$-axis. For this reason one could call these chiral ladder operators since they relate directly to quanta with a definite handedness.
Introducing now the dimensionless coordinate $z=\sqrt{\lambda}(x+i y)$ and its complex conjugate $z^{*}$ where $\lambda=\mu \omega / \hbar$ as above, we can write these new ladder operators in the coordinate representation as

$$
\begin{equation*}
a_{+}=\frac{z^{*}}{2}+\frac{\partial}{\partial z}, \quad a_{-}=\frac{z}{2}+\frac{\partial}{\partial z^{*}}, \tag{5.36}
\end{equation*}
$$

Since the ground state wave function of the oscillator $\psi_{0,0}$ is defined by $a_{ \pm} \psi_{0,0}=0$, it will have be

$$
\begin{equation*}
\psi_{0,0}=\sqrt{\frac{1}{\pi}} e^{-z^{*} z / 2} \tag{5.37}
\end{equation*}
$$

where the factor in front is a normalization factor. This is just the product of the ground-state wave functions in the $x$ and $y$ directions respectively. Since we can write the circular raising operators as

$$
a_{+}^{\dagger}=-e^{z^{*} z / 2} \frac{\partial}{\partial z^{*}} e^{-z^{*} z / 2}, \quad a_{-}^{\dagger}=-e^{z^{*} z / 2} \frac{\partial}{\partial z} e^{-z^{*} z / 2}
$$

we now find the excited state with angular momentum $+\hbar m$ where the quantum number $m$ is positive, to be

$$
\begin{equation*}
\psi_{m, 0}=\sqrt{\frac{1}{\pi}} \frac{a_{+}^{\dagger m}}{\sqrt{m!}} e^{-z^{*} z / 2}=\sqrt{\frac{1}{\pi}} \frac{z^{m}}{\sqrt{m!}} e^{-z^{*} z / 2} \tag{5.38}
\end{equation*}
$$

Using instead the operator $a_{-}^{\dagger}$ we find the state with angular momentum $-\hbar m$ where $z$ will be replaced by $z^{*}$ in the wave function.
Radial exitations can now be obtained by the operator $\widehat{b}^{\dagger}=\widehat{a}_{+}^{\dagger} \widehat{a}_{-}^{\dagger}$ which increase the energy with $2 \hbar \omega$ but does not change the angular momentum. A general excited state will therefore be

$$
\begin{align*}
\psi_{m, n}\left(z, z^{*}\right) & =\sqrt{\frac{m!}{n!(m+n)!}}\left(b^{\dagger}\right)^{n} \psi_{m, 0} \\
& =\frac{(-1)^{n}}{\sqrt{\pi n!(m+n)!}} e^{z^{*} z / 2}\left(\frac{\partial}{\partial z}\right)^{n}\left(z^{m+n} e^{-z^{*} z}\right) \tag{5.39}
\end{align*}
$$

This should now be compared with the solution (5.32). In fact, they are seen to agree when we introduce the scalar variable $t=z^{*} z=\lambda r^{2}$ so that we can write our newly constructed eigenfunction as

$$
\begin{equation*}
\psi_{m, n}=\frac{(-1)^{n}}{\sqrt{\pi n!(m+n)!}} e^{-z^{*} z / 2} z^{m} e^{t} t^{-m}\left(\frac{d}{d t}\right)^{n}\left(t^{m+n} e^{-t}\right) \tag{5.40}
\end{equation*}
$$

Comparing with the definition of the Laguerre polynomials in (5.28), we thus have

$$
\begin{equation*}
\psi_{m, n}\left(z, z^{*}\right)=(-1)^{n} \sqrt{\frac{n!}{\pi(m+n)!}} e^{-z^{*} z / 2} z^{m} L_{n}^{m}\left(z^{*} z\right) \tag{5.41}
\end{equation*}
$$

which is just our previous solution (5.32), but now with even correct normalization.

### 5.5 Spherical rotations and the generators

Let us first consider a particle in the $x y$-plane at position $\mathbf{r}=x \mathbf{e}_{x}+y \mathbf{e}_{y}$ as shown in the figure. It has coordinates $x=r \cos \phi$ and $y=r \sin \phi$. Under a rotation by an angle $\alpha$ around the $z$-axis, the vector changes into the new vector $\mathbf{r}^{\prime}=x^{\prime} \mathbf{e}_{x}+y^{\prime} \mathbf{e}_{y}$ with components

$$
x^{\prime}=r \cos (\alpha+\phi)=r(\cos \alpha \cos \phi-\sin \alpha \sin \phi)=x \cos \alpha-y \sin \alpha
$$



Figure 5.3: Rotation by an angle $\phi$ of a two-dimensional vector $\mathbf{r}$ into the final vector $\mathbf{r}^{\prime}$.
and

$$
y^{\prime}=r \sin (\alpha+\phi)=r(\sin \alpha \cos \phi+\cos \alpha \sin \phi)=x \sin \alpha+y \cos \alpha
$$

which becomes on matrix form

$$
\binom{x^{\prime}}{y^{\prime}}=\left(\begin{array}{rr}
\cos \alpha & -\sin \alpha  \tag{5.42}\\
\sin \alpha & \cos \alpha
\end{array}\right)\binom{x}{y}
$$

For general rotation $\mathbf{r} \rightarrow \mathbf{r}^{\prime}=R \mathbf{r}$ around an arbitrary axis the matrix $R$ would be $3 \times 3$ and would be parametrized by three rotation angles. These could for instance be the three Euler angles $\alpha, \beta$ and $\gamma$ used in classical mechanics.

Returning now to the 2-dimensional rotation (5.42) in the infinitesemal limit where $\alpha \rightarrow \delta \alpha$ we thus have $x^{\prime}=x-y \delta \alpha$ and $y^{\prime}=y+x \delta \alpha$ when we neglect higher order terms. The change in the original vector $\mathbf{r}$ is therefore be written as $\delta \mathbf{r}=\mathbf{r}^{\prime}-\mathbf{r}=\delta \alpha \mathbf{e}_{z} \wedge \mathbf{r}$. Introducing now the vector $\delta \boldsymbol{\alpha}=\delta \alpha \mathbf{e}_{z}$ which is along the rotation axis and has a magnitude equal to the rotation angle, we have the compact result

$$
\begin{equation*}
\delta \mathbf{r}=\delta \boldsymbol{\alpha} \wedge \mathbf{r} \tag{5.43}
\end{equation*}
$$

A little geometrical thinking will convince us that this expression is valid for an infinitesemal rotation around any axis $\mathbf{n}$, i.e. $\delta \boldsymbol{\alpha}=\mathbf{n} \delta \alpha$ when $\mathbf{n}$ is a unit vector. Similarly, a finite rotation by the angle $\alpha$ around the same axis is given by the rotation vector $\boldsymbol{\alpha}=\mathbf{n} \alpha$.

In a quantum description a particle at position $\mathbf{r}$ is in the position eigenstate $|\mathbf{r}\rangle$. Under a rotation by a finite angle $\phi$ it changes into

$$
\begin{equation*}
\left|\mathbf{r}^{\prime}\right\rangle=\widehat{R}(\boldsymbol{\alpha})|\mathbf{r}\rangle \tag{5.44}
\end{equation*}
$$

where $\widehat{R}$ is the rotation operator. Since the norm of this ket vector is conserved under this transformation, it is unitary in the sense that

$$
\begin{equation*}
\widehat{R}^{\dagger}(\boldsymbol{\alpha})=\widehat{R}^{-1}(\boldsymbol{\alpha})=\widehat{R}(-\boldsymbol{\alpha}) \tag{5.45}
\end{equation*}
$$

as the corresponding translation operator was in Chapter 3.
Again we can build up a finite rotation by combining many infinitesemal ones of the form (5.9) in the plane. We thus write

$$
\begin{equation*}
\widehat{R}(\delta \boldsymbol{\alpha})=\widehat{I}-\frac{i}{\hbar} \delta \boldsymbol{\alpha} \cdot \widehat{\mathbf{L}} \tag{5.46}
\end{equation*}
$$

where the three Hermitean generators $\widehat{\mathbf{L}}=\left(\widehat{L}_{x}, \widehat{L}_{y}, \widehat{L}_{z}\right)$ will be shown to be the threedimensional angular momentum operators. Iterating many such infinitesemal rotations, we thus can write for a finite rotation

$$
\begin{equation*}
\widehat{R}(\boldsymbol{\alpha})=e^{-i \boldsymbol{\alpha} \cdot \widehat{\mathbf{L}} / \hbar} \tag{5.47}
\end{equation*}
$$

for a general rotation $\boldsymbol{\alpha}=\left(\alpha_{x}, \alpha_{y}, \alpha_{z}\right)$. Writing instead $\boldsymbol{\alpha}=\mathbf{n} \alpha$ where $\mathbf{n}$ is the rotation axis and $\alpha$ is the amount of rotation, we can equivalently denote the same rotation by

$$
\begin{equation*}
\widehat{R}_{\mathbf{n}}(\alpha)=e^{-i \alpha \mathbf{n} \cdot \widehat{\mathbf{L}} / \hbar} \tag{5.48}
\end{equation*}
$$

It is generated by the differential operator

$$
\begin{equation*}
\mathbf{n} \cdot \mathbf{L}=-i \hbar \frac{\partial}{\partial \alpha} \tag{5.49}
\end{equation*}
$$

in complete analogy with (5.13).
Rotations by the same angle around different axes are said to be similar or equivalent. Let us again consider the general rotation (5.48). This can now be used to find the operator $\widehat{R}_{\mathbf{n}^{\prime}}(\alpha)$ for the rotation about another axis $\mathbf{n}^{\prime}$ by the same angle $\alpha$. For that we need the rotation operator $\widehat{S}_{\mathbf{n} \rightarrow \mathbf{n}^{\prime}}$ which takes the unit vector $\mathbf{n}$ into the new rotation axis $\mathbf{n}^{\prime}$. From the resulting geometry of this operation, we see that we then have the equality

$$
\begin{equation*}
\widehat{R}_{\mathbf{n}^{\prime}}(\alpha)=\widehat{S}_{\mathbf{n} \rightarrow \mathbf{n}^{\prime}} \widehat{R}_{\mathbf{n}}(\alpha) \widehat{S}_{\mathbf{n} \rightarrow \mathbf{n}^{\prime}}^{\dagger} \tag{5.50}
\end{equation*}
$$

The mathematical name of this equivalence is a similarity transformation.
Just as the position eigenvectors transform as (5.44) under rotations, any other general state vector $|\psi\rangle$ is defined to transform the same way, i.e.

$$
\begin{equation*}
|\psi\rangle \rightarrow\left|\psi^{\prime}\right\rangle=\widehat{R}(\boldsymbol{\alpha})|\psi\rangle \tag{5.51}
\end{equation*}
$$

The corresponding wave function $\psi(\mathbf{r})=\langle\mathbf{r} \mid \psi\rangle$ will thus change into the transformed wave function $\psi^{\prime}(\mathbf{r})=\left\langle\mathbf{r} \mid \psi^{\prime}\right\rangle$. Its dependence on the original function follows from

$$
\begin{equation*}
\psi(\mathbf{r})=\langle\mathbf{r} \mid \psi\rangle=\langle\mathbf{r}| \widehat{R}^{\dagger}(\boldsymbol{\alpha}) \widehat{R}(\boldsymbol{\alpha})|\psi\rangle=\left\langle\mathbf{r}^{\prime} \mid \psi^{\prime}\right\rangle=\psi^{\prime}\left(\mathbf{r}^{\prime}\right) \tag{5.52}
\end{equation*}
$$

It is thus said to transform as a scalar function where the new function at the new point equals the original function at the original point. The change in the function under an infinitesemal rotation as measured at the same point is therefore

$$
\begin{aligned}
\delta \psi & =\psi^{\prime}(\mathbf{r})-\psi(\mathbf{r})=\psi(\mathbf{r}-\delta \mathbf{r})-\psi(\mathbf{r})=-\delta \mathbf{r} \cdot \boldsymbol{\nabla} \psi \\
& =-(\delta \boldsymbol{\alpha} \wedge \mathbf{r}) \cdot \boldsymbol{\nabla} \psi=-\delta \boldsymbol{\alpha} \cdot(\mathbf{r} \wedge \boldsymbol{\nabla} \psi)=-\frac{i}{\hbar} \delta \boldsymbol{\alpha} \cdot \mathbf{L} \psi
\end{aligned}
$$

Here we have written the operator $\mathbf{L}=-i \hbar \mathbf{r} \wedge \boldsymbol{\nabla}=\mathbf{r} \wedge \mathbf{p}$ in the coordinate representation where the momentum operator $\widehat{\mathbf{p}}$ is represented by the differential operator $\mathbf{p}=-i \hbar \boldsymbol{\nabla}$. But at the same time we have from (5.46) that this change in the wave function is given by

$$
\delta \psi=\langle\mathbf{r}| \widehat{R}(\delta \boldsymbol{\alpha})|\psi\rangle-\langle\mathbf{r} \mid \psi\rangle=-\frac{i}{\hbar} \delta \boldsymbol{\alpha} \cdot\langle\mathbf{r}| \widehat{\mathbf{L}}|\psi\rangle
$$

Comparing these two expressions, we therefore find that

$$
\langle\mathbf{r}| \widehat{\mathbf{L}}|\psi\rangle=-i \hbar \mathbf{r} \wedge \nabla\langle\mathbf{r} \mid \psi\rangle
$$

which proves that the rotation generators are just the quantized angular momentum operator $\widehat{\mathbf{L}}=\widehat{\mathbf{r}} \wedge \widehat{\mathbf{p}}$.

Since the operators $\widehat{\mathbf{r}}$ and $\widehat{\mathbf{p}}$ don't commute in general, so will also the different components $\widehat{L}_{i}$ of the angular momentum operator not commute with each other in general. In order to calculate these non-zero commutators, it is very convenient to introduce Einstein's summation convention. It says that in any expression where there are two equal indices, they should be summed over. For instance, the scalar product of two vectors $\mathbf{A}$ and $\mathbf{B}$ we write simply as $\mathbf{A} \cdot \mathbf{B}=A_{i} B_{i}=A_{k} B_{k}$ when expressed by their Cartesian components. The vector $\mathbf{A}$ can therefore be written as $\mathbf{A}=A_{i} \mathbf{e}_{i}$ where the unit basis vectors satisfy $\mathbf{e}_{i} \cdot \mathbf{e}_{j}=\delta_{i j}$ when expressed in terms of the Kroenecker $\delta$-symbol. More detailed, we then have $\mathbf{A} \cdot \mathbf{B}=A_{i} B_{j} \delta_{i j}=A_{i} B_{i}$.
We can similarly express the vector or wedge product between two vectors $\mathbf{A}$ and $\mathbf{B}$ as

$$
\begin{equation*}
\mathbf{C}=\mathbf{A} \wedge \mathbf{B}=\epsilon_{i j k} A_{i} B_{j} \mathbf{e}_{k} \tag{5.53}
\end{equation*}
$$

when we make use of the $\epsilon$-symbol of Levi-Civita. It is defined to be antisymmetric under interchange of any two indices. It is therefore zero when any two indices are equal. The only non-zero values then follow from $\epsilon_{123}=+1$ which gives $\epsilon_{312}=\epsilon_{231}=1$ while $\epsilon_{132}=\epsilon_{321}=-1$ and so on. For example, the third component of the above vector $\mathbf{C}$ is then $C_{3}=\epsilon_{i j 3} A_{i} B_{j}=A_{1} B_{2}-A_{2} B_{1}$ making use of the summation convention. Needless to say, this is the correct result.

With these new tools, we can now write the components of the angular momentum operator as $\widehat{L}_{i}=\epsilon_{i k m} \widehat{x}_{k} \widehat{p}_{m}$ and we can start to calculate. First we consider

$$
\left[\widehat{L}_{i}, \widehat{x}_{j}\right]=\epsilon_{i k m}\left[\widehat{x}_{k} \widehat{p}_{m}, \widehat{x}_{j}\right]=\epsilon_{i k m}\left(\widehat{x}_{k}\left[\widehat{p}_{m}, \widehat{x}_{j}\right]+\left[\widehat{x}_{k}, \widehat{x}_{j}\right] \widehat{p}_{m}\right)
$$

Since the last commutator is zero and $\left[\widehat{p}_{m}, \widehat{x}_{j}\right]=-i \hbar \delta_{m j}$, we get

$$
\begin{equation*}
\left[\widehat{L}_{i}, \widehat{x}_{j}\right]=-i \hbar \epsilon_{i k m} \delta_{m j} \widehat{x}_{k}=i \hbar \epsilon_{i j k} \widehat{x}_{k} \tag{5.54}
\end{equation*}
$$

In the same way we find $\left[\widehat{L}_{i}, \widehat{x}_{j}\right]=i \hbar \epsilon_{i j m} \widehat{p}_{m}$. Since both these vector operators transform the same way, we might think that vector operator transforms this way and in particular the angular momentum operator it self. We thus expect

$$
\begin{equation*}
\left[\widehat{L}_{i}, \widehat{L}_{j}\right]=i \hbar \epsilon_{i j k} \widehat{L}_{k} \tag{5.55}
\end{equation*}
$$

This is easy to verify. For instance, making use of the above commutators, we find

$$
\begin{equation*}
\left[\widehat{L}_{x}, \widehat{L}_{y}\right]=\left[\widehat{L}_{x}, \widehat{z} \widehat{p}_{x}-\widehat{x} \widehat{p}_{z}\right]=\left[\widehat{L}_{x}, \widehat{z}\right] \widehat{p}_{x}-\widehat{x}\left[\widehat{L}_{x}, \widehat{p}_{z}\right]=-i \hbar \widehat{y} \widehat{p}_{x}+i \hbar \widehat{x} \widehat{p}_{y}=i \hbar \widehat{L}_{z} \tag{5.56}
\end{equation*}
$$

and correspondingly for other components.
The basic angular momentum commutator (5.55) is sometimes written other ways in the literature. For instance, introducing the angular momentum along the unit vector a defined as $\widehat{L}_{\mathbf{a}}=\mathbf{a} \cdot \widehat{\mathbf{L}}=a_{i} \widehat{L}_{i}$, we find from (5.55) by multiplying both sides by $a_{i}$ and the components $b_{i}$ of another unit vector $\mathbf{b}$,

$$
\begin{equation*}
\left[\widehat{L}_{\mathbf{a}}, \widehat{L}_{\mathbf{b}}\right]=i \hbar \widehat{L}_{\mathbf{a} \wedge \mathbf{b}} \tag{5.57}
\end{equation*}
$$

using the expression for the wedge product written in terms of the Levi-Civita symbol. Equivalently, since $(\widehat{\mathbf{L}} \wedge \widehat{\mathbf{L}})_{k}=\epsilon_{i j k} \widehat{L}_{i} \widehat{L}_{j}=i \hbar \widehat{L}_{k}$, we also have

$$
\begin{equation*}
\widehat{\mathbf{L}} \wedge \widehat{\mathbf{L}}=i \hbar \widehat{\mathbf{L}} \tag{5.58}
\end{equation*}
$$

We will later meet other non-zero wedge products between a vector and itself when the components don't commute with each other.

### 5.6 Quantization of spin

Since the three components of the angular momentum operator don't commute with each other, we cannot find eigenstates of angular momentum which are simultaneously eigenstates for all three components. However, since the total angular momentum $\widehat{\mathbf{L}}^{2}$ commutes with the components,

$$
\begin{align*}
{\left[\widehat{\mathbf{L}}^{2}, \widehat{L}_{j}\right] } & =\left[\widehat{L}_{i} \widehat{L}_{i}, \widehat{L}_{j}\right]=\widehat{L}_{i}\left[\widehat{L}_{i}, \widehat{L}_{j}\right]+\left[\widehat{L}_{i}, \widehat{L}_{j}\right] \widehat{L}_{i}  \tag{5.59}\\
& =i \hbar \epsilon_{i j k}\left(\widehat{L}_{i} \widehat{L}_{k}+\widehat{L}_{k} \widehat{L}_{i}\right)=0 \tag{5.60}
\end{align*}
$$

so its eigenvalues can be determined together with the eigenvalues of one of the components. It is customary to take this component to be $\widehat{L}_{z}$. Denoting these eigenvalues with $\lambda$ and $\mu$, we will then have eigenstates $|\lambda, \mu\rangle$ satisfying

$$
\widehat{\mathbf{L}}^{2}|\lambda, \mu\rangle=\lambda|\lambda, \mu\rangle, \quad \widehat{L}_{z}|\lambda, \mu\rangle=\mu|\lambda, \mu\rangle
$$

Since $\widehat{\mathbf{L}}^{2}=\widehat{L}_{x}^{2}+\widehat{L}_{y}^{2}+\widehat{L}_{z}^{2}$ and the operator $\widehat{L}_{x}^{2}+\widehat{L}_{y}^{2}$ is always positive or zero, we must have the condition $\lambda \geq \mu^{2}$ always satisfied bewteen these eigenvalues.
It is now very convenient to replace the two remaining component operators $\widehat{L}_{x}$ and $\widehat{L}_{y}$ with the linear combinations

$$
\begin{equation*}
\widehat{L}_{ \pm}=\widehat{L}_{x} \pm i \widehat{L}_{y} \tag{5.61}
\end{equation*}
$$

These are no longer Hermitian, but $\widehat{L}_{ \pm}^{\dagger}=\widehat{L}_{\mp}$. It then follows that the product $\widehat{L}_{+} \widehat{L}_{-}=$ $\widehat{L}_{x}^{2}+\widehat{L}_{y}^{2}+i\left[\widehat{L}_{x}, \widehat{L}_{y}\right]$ or

$$
\begin{equation*}
\widehat{L}_{+} \widehat{L}_{-}=\widehat{\mathbf{L}}^{2}-\widehat{L}_{z}^{2}+\hbar \widehat{L}_{z} \tag{5.62}
\end{equation*}
$$

and similarly

$$
\begin{equation*}
\widehat{L}_{-} \widehat{L}_{+}=\widehat{\mathbf{L}}^{2}-\widehat{L}_{z}^{2}-\hbar \widehat{L}_{z} \tag{5.63}
\end{equation*}
$$

By subtraction we then have

$$
\begin{equation*}
\left[\widehat{L}_{+}, \widehat{L}_{-}\right]=2 \hbar \widehat{L}_{z} \tag{5.64}
\end{equation*}
$$

The commutators between $\widehat{L}_{z}$ and the two other components, are now replaced by

$$
\begin{equation*}
\left[\widehat{L}_{z}, \widehat{L}_{ \pm}\right]= \pm \hbar \widehat{L}_{ \pm} \tag{5.65}
\end{equation*}
$$

It thus looks like $\widehat{L}_{ \pm}$are raising and lowering operators for the eigenvalues of $\widehat{L}_{z}$. And indeed, consider the above eigenstate $|\lambda, \mu\rangle$ multiplied by $\widehat{L}_{+}$. It is a new eigenstate since

$$
\widehat{L}_{z} \widehat{L}_{+}|\lambda, \mu\rangle=\left(\widehat{L}_{+} \widehat{L}_{z}+\hbar \widehat{L}_{+}\right)|\lambda, \mu\rangle=(\mu+\hbar) \widehat{L}_{+}|\lambda, \mu\rangle
$$

and the eigenvalue has increased with $+\hbar$. It is therefore natural to measure the eigenvalue $\mu$ in units of $\hbar$ and write $\mu=\hbar m$ where the quantum number $m$ is so far unknown. In the same way $\widehat{L}_{-}$is found to lower the eigenvalue by $\hbar$. We therefore have

$$
\widehat{L}_{ \pm}|\lambda, m\rangle=|\lambda, m \pm 1\rangle
$$

when we ignore normalization constants. Since the commutator (5.60) is zero, the state $\widehat{L}_{ \pm}|\lambda, m\rangle$ is still an eigenstate of $\widehat{\mathbf{L}}^{2}$ with the same eigenvalue $\lambda$.
When we now continue to apply the raising operator $\widehat{L}_{+}$to the state $|\lambda, m\rangle$ the eigenvalues of $\widehat{L}_{z}$ will increase and eventually the above bound $\lambda \geq \hbar^{2} m^{2}$ will be violated since $\lambda$ doesn't change. There must therefore be an upper state $|\lambda, l\rangle$ so that $\widehat{L}_{+}|\lambda, l\rangle=0$. This state is often also called the highest state. Applying now the lowering operator to it and using the product(5.63), one obtains

$$
\widehat{L}_{-} \widehat{L}_{+}|\lambda, l\rangle=0=\left(\widehat{\mathbf{L}}^{2}-\widehat{L}_{z}^{2}-\hbar \widehat{L}_{z}\right)|\lambda, l\rangle=\left(\lambda-\hbar^{2} l^{2}-\hbar^{2} l\right)|\lambda, l\rangle
$$

and therefore the important result $\lambda=\hbar^{2} l(l+1)$.
So far $l$ is just some real number. It can now be determined by applying the lowering operator $\widehat{L}_{-}$to this highest state. Each time the eigenvalue $m$ is decreased by 1 and eventually we obtain a state $|\lambda, l-n\rangle$ where $n$ is a positive integer or zero. This lowest state must now obey $\widehat{L}_{-}|\lambda, l-n\rangle=0$ in order for the same bound not to be violated. Multiplying this condition by $\widehat{L}_{+}$and now using the product (5.62), we have

$$
\begin{aligned}
\widehat{L}_{+} \widehat{L}_{-}|\lambda, l-n\rangle=0 & =\left(\widehat{\mathbf{L}}^{2}-\widehat{L}_{z}^{2}+\hbar \widehat{L}_{z}\right)|\lambda, l-n\rangle \\
& =\left[\lambda-\hbar^{2}(l-n)^{2}-\hbar^{2}(l-n)\right]|\lambda, l-n\rangle
\end{aligned}
$$

Since we already have found that $\lambda=\hbar^{2} l(l+1)$, it follows that

$$
l(l+1)-(l-n)^{2}+(l-n)=0
$$

or $2 l(n+1)=n(n+1)$ which gives $l=n / 2$. The magnitude of the angular momentum is therefore quantized and given by the possible values $l=0,1 / 2,1,3 / 2 \cdots$ of the corresponding quantum number.

We can therefore sum up our results for the quantization of angular momentum by the two eigenvalue equations

$$
\begin{align*}
\widehat{\mathbf{L}}^{2}|l, m\rangle & =\hbar^{2} l(l+1)|l, m\rangle  \tag{5.66}\\
\widehat{L}_{z}|l, m\rangle & =\hbar m|l, m\rangle \tag{5.67}
\end{align*}
$$

where the quantum number $l$ is an integer or half integer while the other quantum number can take on one of the values $m=(l, l-1, l-2, \cdots,-l+1,-l)$ corresponding to $2 l+1$ possible directions of the quantized angular momentum with respect to the $z$-axis. It is very practical to have these eigenstates orthonormalized in the sense that $\left\langle l, m \mid l^{\prime}, m^{\prime}\right\rangle=\delta_{l l^{\prime}} \delta_{m m^{\prime}}$. We must thus have a relation like $\widehat{L}_{+}|l, m\rangle=N_{l m}|l, m+1\rangle$ with some unknown coefficient $N_{l m}$. The normalization condition then requires that $\left|N_{l m}\right|^{2}=\langle l, m| \widehat{L}_{-} \widehat{L}_{+}|l, m\rangle$. Using now the result (5.63) for the product $\widehat{L}_{-} \widehat{L}_{+}$, it follows that $\left|N_{l m}\right|^{2}=\hbar^{2}\left[l(l+1)-m^{2}-m\right]$ and therefore

$$
\begin{equation*}
\widehat{L}_{+}|l, m\rangle=\hbar \sqrt{l(l+1)-m(m+1)}|l, m+1\rangle \tag{5.68}
\end{equation*}
$$

when we set a possible phase factor equal to one. Similarly one finds

$$
\begin{equation*}
\widehat{L}_{-}|l, m\rangle=\hbar \sqrt{l(l+1)-m(m-1)}|l, m-1\rangle \tag{5.69}
\end{equation*}
$$

These are very useful in practical calculations. A simple way to remember both, is to combine them in the single expression

$$
\begin{equation*}
\widehat{L}_{ \pm}|l, m\rangle=\hbar \sqrt{\left(l+m_{>}\right)\left(l-m_{<}\right)}|l, m \pm 1\rangle \tag{5.70}
\end{equation*}
$$

where $m_{>}$is the largest of the two $m$-values appearing in the expression while $m_{<}$is the smallest of the two. The largest appears with a plus sign and the smallest with minus.
Within this basis of eigenstates corresponding to a value $l$ of the total quantum number, we can now easily find matrix representations of the angular momentum operators defined by

$$
\begin{equation*}
\left(L_{i}\right)_{m m^{\prime}}=\langle l, m| \widehat{L}_{i}\left|l, m^{\prime}\right\rangle \tag{5.71}
\end{equation*}
$$

These will be $(2 l+1) \times(2 l+1)$ matrices satifying the fundamental commutator $\left[L_{i}, L_{j}\right]=$ $i \hbar \epsilon_{i j k} L_{k}$. We have derived it using the commutators of the orbital angular momentum operators $\widehat{\mathbf{L}}=\widehat{\mathbf{r}} \wedge \widehat{\mathbf{p}}$. But we will in the following generalize the above results to any set of three operators having the same commutators. In this situation it is more appropriate to call these operators for general spin operators.

### 5.7 Spin-1/2 and the Pauli matrices

The smallest, non-trivial spin value is obtained for $l=1 / 2$. This is a special and very important case and it is customary to denote the corresponding spin operators by $\widehat{\mathbf{S}}$ with $\left[\widehat{S}_{x}, \widehat{S}_{y}\right]=i \hbar \widehat{S}_{z}$. Corresponding to (5.66) and (5.67), we then have the eigenvalue equations

$$
\begin{align*}
\widehat{\mathbf{S}}^{2}|1 / 2, m\rangle & =(3 / 4) \hbar^{2}|1 / 2, m\rangle  \tag{5.72}\\
\widehat{S}_{z}|1 / 2, m\rangle & =m \hbar|1 / 2, m\rangle \tag{5.73}
\end{align*}
$$

where now the second quantum number can only take the two values $m= \pm 1 / 2$. Since the basis vectors $|1 / 2, \pm 1 / 2\rangle$ are eigenstates for $\widehat{S}_{z}$, the corresponding matrix $S_{z}$ is diagonal,

$$
S_{z}=\hbar\left(\begin{array}{cc}
1 / 2 & 0  \tag{5.74}\\
0 & -1 / 2
\end{array}\right)=\frac{1}{2} \hbar \sigma_{z}
$$

where

$$
\sigma_{z}=\left(\begin{array}{rr}
1 & 0  \tag{5.75}\\
0 & -1
\end{array}\right)
$$

The two other matrices $S_{x}$ and $S_{y}$ follow from (5.70) which yield only two non-zero matrix elements from

$$
\widehat{S}_{+}|1 / 2,-1 / 2\rangle=\hbar|1 / 2,+1 / 2\rangle, \quad \widehat{S}_{-}|1 / 2,+1 / 2\rangle=\hbar|1 / 2,-1 / 2\rangle
$$

corresponding to

$$
S_{+}=\hbar\left(\begin{array}{ll}
0 & 1  \tag{5.76}\\
0 & 0
\end{array}\right), \quad S_{-}=\hbar\left(\begin{array}{ll}
0 & 0 \\
1 & 0
\end{array}\right)
$$

We can therefore write the two Cartesian components $S_{x}=\left(S_{+}+S_{-}\right) / 2=(\hbar / 2) \sigma_{x}$ and $S_{y}=\left(S_{+}-S_{-}\right) / 2 i=(\hbar / 2) \sigma_{y}$ in terms of the two fundamental matrices

$$
\sigma_{x}=\left(\begin{array}{cc}
0 & 1  \tag{5.77}\\
1 & 0
\end{array}\right), \quad \sigma_{y}=\left(\begin{array}{rr}
0 & -i \\
i & 0
\end{array}\right)
$$

All the representation matrices for spin $S=1 / 2$ thus form the vector

$$
\begin{equation*}
\mathbf{S}=\frac{1}{2} \hbar \boldsymbol{\sigma} \tag{5.78}
\end{equation*}
$$

where the components of the vector $\boldsymbol{\sigma}=\left(\sigma_{x}, \sigma_{y}, \sigma_{z}\right)$ are the three Pauli matrices. Since the spin operator $\widehat{\mathbf{S}}$ is self-adjoint, they are Hermitean matrices.
In terms of these fundamental matrices the basic commutator $\left[\widehat{S}_{i}, \widehat{S}_{j}\right]=i \hbar \epsilon_{i j k} \widehat{S}_{k}$ will be represented by the matrix commutator

$$
\begin{equation*}
\left[\sigma_{i}, \sigma_{j}\right]=2 i \epsilon_{i j k} \sigma_{k} \tag{5.79}
\end{equation*}
$$

or $\left[\sigma_{x}, \sigma_{y}\right]=2 i \sigma_{z}$ which is easily seen to be satisfied. Indeed, we find

$$
\left(\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right)\left(\begin{array}{rr}
0 & -i \\
i & 0
\end{array}\right)=\left(\begin{array}{rr}
i & 0 \\
0 & -i
\end{array}\right)
$$

or $\sigma_{x} \sigma_{y}=i \sigma_{z}$ and similarly $\sigma_{y} \sigma_{x}=-i \sigma_{z}$. This means that $\sigma_{x}$ and $\sigma_{y}$ anticommute in the sense that $\sigma_{x} \sigma_{y}+\sigma_{y} \sigma_{x}=0$. One easily verifies that this is the case for any product of two different Pauli matrices. On the other hand, the product of two equal Pauli matrices is always given by the unit matrix, $\sigma_{x}^{2}=\sigma_{y}^{2}=\sigma_{z}^{2}=I$, which we usually write equal to one. The general anticommutator of two Pauli matrices is therefore

$$
\begin{equation*}
\sigma_{i} \sigma_{j}+\sigma_{j} \sigma_{i}=2 \delta_{i j} \tag{5.80}
\end{equation*}
$$

Combining this with the commutator in (5.79), we have

$$
\begin{equation*}
\sigma_{i} \sigma_{j}=\delta_{i j}+i \epsilon_{i j k} \sigma_{k} \tag{5.81}
\end{equation*}
$$

It should be stressed that these relations are only valid for the representation matrices of the spin- $1 / 2$ operators and don't hold for other spins.

If we now are given a vector $\mathbf{a}=\left(a_{x}, a_{y}, a_{z}\right)$, we can call the product matrix

$$
\boldsymbol{\sigma} \cdot \mathbf{a}=\sigma_{x} a_{x}+\sigma_{y} a_{y}+\sigma_{z} a_{z}=\left(\begin{array}{cc}
a_{z} & a_{-} \\
a_{+} & -a_{z}
\end{array}\right)
$$

where $a_{ \pm}=a_{x} \pm i a_{y}$, for the spin along this vector. Squaring it, one finds

$$
\left(\begin{array}{cc}
a_{z} & a_{-} \\
a_{+} & -a_{z}
\end{array}\right)\left(\begin{array}{cc}
a_{z} & a_{-} \\
a_{+} & -a_{z}
\end{array}\right)=\left(\begin{array}{cc}
a_{x}^{2}+a_{y}^{2}+a_{z}^{2} & 0 \\
0 & a_{x}^{2}+a_{y}^{2}+a_{z}^{2}
\end{array}\right)=\mathbf{a} \cdot \mathbf{a}
$$

or $(\boldsymbol{\sigma} \cdot \mathbf{a})^{2}=\mathbf{a} \cdot \mathbf{a}$. This important result also follows directly as a special case from the general product (5.81). Given a second vector $\mathbf{b}=\left(b_{x}, b_{y}, b_{z}\right)$, we multiply both sides by $a_{i} b_{j}$. Now using again the Einstein summation convention and the definition of the vector product in terms of the Levi-Civita symbol, we get immediately

$$
\begin{equation*}
(\boldsymbol{\sigma} \cdot \mathbf{a})(\boldsymbol{\sigma} \cdot \mathbf{b})=\mathbf{a} \cdot \mathbf{b}+i \boldsymbol{\sigma} \cdot(\mathbf{a} \wedge \mathbf{b}) \tag{5.82}
\end{equation*}
$$

In addition to its elegance, this beautiful formula has many applications as we will later see.

Since this spin representation space is two-dimensional, the corresponding state vectors will be two-component column matrices

$$
\begin{equation*}
\chi=\binom{\chi_{1}}{\chi_{2}} \tag{5.83}
\end{equation*}
$$

called Pauli spinors. The components $\chi_{1}$ and $\chi_{2}$ are in general complex numbers so that the norm becomes

$$
\begin{equation*}
\chi^{\dagger} \chi=\chi_{1}^{*} \chi_{1}+\chi_{2}^{*} \chi_{2} \tag{5.84}
\end{equation*}
$$

We can write the general Pauli spinor (5.83) as $\chi=\chi_{1} \alpha+\chi_{2} \beta$ where the unit-norm spinors

$$
\begin{equation*}
\alpha=\binom{1}{0}, \quad \beta=\binom{0}{1} \tag{5.85}
\end{equation*}
$$

provide an orthogonal basis in this space. These are seen to represent the basis ket vectors as $|1 / 2,+1 / 2\rangle \rightarrow \alpha$ and $|1 / 2,-1 / 2\rangle \rightarrow \beta$. They are therefore eigenspinors of $\sigma_{z}$ as is also verified by $\sigma_{z} \alpha=+\alpha$ and $\sigma_{z} \beta=-\beta$.

A general rotation of a spinor $\chi$ is now given by the matrix version of (5.51), i.e.

$$
\begin{equation*}
\chi \rightarrow \chi^{\prime}=R(\boldsymbol{\alpha}) \chi \tag{5.86}
\end{equation*}
$$

where now $R$ is a $2 \times 2$ rotation matrix

$$
\begin{equation*}
R(\boldsymbol{\alpha})=e^{-i \boldsymbol{\alpha} \cdot \mathbf{S} / \hbar}=e^{-i \boldsymbol{\alpha} \cdot \boldsymbol{\sigma} / 2} \tag{5.87}
\end{equation*}
$$

Introducing the unit vector $\mathbf{n}$ for the direction of rotation, we then have

$$
\begin{aligned}
R_{\mathbf{n}}(\alpha) & =e^{-i \alpha \boldsymbol{\sigma} \cdot \mathbf{n} / 2} \\
& =1-i \frac{\alpha}{2} \boldsymbol{\sigma} \cdot \mathbf{n}+\frac{1}{2!}\left(-i \frac{\alpha}{2} \boldsymbol{\sigma} \cdot \mathbf{n}\right)^{2}+\frac{1}{3!}\left(-i \frac{\alpha}{2} \boldsymbol{\sigma} \cdot \mathbf{n}\right)^{3}+\ldots
\end{aligned}
$$

Now using $(\boldsymbol{\sigma} \cdot \mathbf{n})^{2}=1$ so that $(\boldsymbol{\sigma} \cdot \mathbf{n})^{3}=\boldsymbol{\sigma} \cdot \mathbf{n}$, one obtains

$$
\begin{align*}
e^{-i \boldsymbol{\alpha} \cdot \boldsymbol{\sigma} / 2} & =\left[1-\frac{1}{2!}\left(\frac{\alpha}{2}\right)^{2}+\frac{1}{4!}\left(\frac{\alpha}{2}\right)^{4}+\ldots\right]-i \boldsymbol{\sigma} \cdot \mathbf{n}\left[\frac{\alpha}{2}-\frac{1}{3!}\left(\frac{\alpha}{2}\right)^{3}+\ldots\right] \\
& =\cos \frac{\alpha}{2}-i \boldsymbol{\sigma} \cdot \mathbf{n} \sin \frac{\alpha}{2} \tag{5.88}
\end{align*}
$$

which can now be used to calculate all rotations of spin- $1 / 2$ spinors.
As a first application, consider the rotation (5.86) by a full turn, i.e. $\alpha=2 \pi$. Since $\sin \pi=0$ and $\cos \pi=-1$, we then find the surprising result $\chi \rightarrow-\chi$. For rotations of ordinary objects we would here expect to get back the original object, i.e. without the minus sign. So in this respect aspin- $1 / 2$ system is a very special object due to its quantum nature. We must make two full rotations in order to get back the original spinor, i.e. for a rotation angle $\alpha=4 \pi$. In many cases this extra minus sign has no physical importance like any other phase factor. But in some special situations it will have observable consequences which are very interesting and important.

A rotation of $\alpha=\pi / 2$ around the $y$-axis is now given by the matrix

$$
R_{y}\left(\frac{\pi}{2}\right)=\cos \frac{\pi}{2}-i \sigma_{y} \sin \frac{\pi}{2}=\sqrt{\frac{1}{2}}\left(\begin{array}{rr}
1 & -1  \tag{5.89}\\
1 & 1
\end{array}\right)
$$

If we then initially have a particle or other system with its spin along the $z$-axis so that it is described by the spin-up spinor $\alpha$, it will after this rotation be in state described by the spinor

$$
\chi^{\prime}=\sqrt{\frac{1}{2}}\left(\begin{array}{rr}
1 & -1 \\
1 & 1
\end{array}\right)\binom{1}{0}=\sqrt{\frac{1}{2}}\binom{1}{1}
$$

But this is an eigenspinor of $\sigma_{x}$ since

$$
\left(\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right)\binom{1}{1}=\binom{1}{1}
$$

and we say that the spin points along the $x$-axis. This is also what we expected from a classical picture.

Under the transformation (5.86) the inner spinor product $\chi^{\dagger} \chi$ remains constant,

$$
\chi^{\dagger} \chi \rightarrow \chi^{\dagger} R^{\dagger} R \chi=\chi^{\dagger} \chi
$$

since the transformation is unitary. But if we instead consider $\chi^{\dagger} \boldsymbol{\sigma} \chi$, it transforms as a 3 -vector. Consider a rotation by the angle $\alpha$ around the $z$-axis. Then $R_{z}(\alpha)=$ $\exp \left(-i \alpha \sigma_{z} / 2\right)$ and therefore

$$
\begin{aligned}
\chi^{\dagger} \sigma_{x} \chi & \rightarrow \chi^{\dagger} e^{i \alpha \sigma_{z} / 2} \sigma_{x} e^{-i \alpha \sigma_{z} / 2} \chi \\
& =\chi^{\dagger}\left[\sigma_{x}+\frac{1}{2}(i \alpha)\left[\sigma_{z}, \sigma_{x}\right]+\frac{1}{8}(i \alpha)^{2}\left[\sigma_{z},\left[\sigma_{z}, \sigma_{x}\right]\right]+\ldots\right] \chi
\end{aligned}
$$

using the Lie formula (2.53). Using now $\left[\sigma_{z}, \sigma_{x}\right]=2 i \sigma_{y}$ and its cyclic cousins, we obtain

$$
\begin{aligned}
\chi^{\dagger} \sigma_{x} \chi & \rightarrow \chi^{\dagger}\left[\sigma_{x}\left(1-\frac{\alpha^{2}}{2!}+\ldots\right)-\sigma_{y}\left(\alpha-\frac{\alpha^{3}}{3!}+\ldots\right)\right] \chi \\
& =\chi^{\dagger} \sigma_{x} \chi \cos \alpha-\chi^{\dagger} \sigma_{y} \chi \sin \alpha
\end{aligned}
$$

In the same way it follows that

$$
\chi^{\dagger} \sigma_{y} \chi \rightarrow \chi^{\dagger} \sigma_{x} \chi \sin \alpha+\chi^{\dagger} \sigma_{y} \chi \cos \alpha
$$

while $\chi^{\dagger} \sigma_{z} \chi$ remains invariant under this particular rotation. Thus the real object $\chi^{\dagger} \boldsymbol{\sigma} \chi$ behaves like a 3 -vector and will return to itself after a full rotation. Such objects are said to belong to the fundamental representation of the ordinary rotation group $S O(3)$. But the 2-dimensional spinor $\chi$ is complex and belongs to the fundamental representation of the complex rotation group $S U(2)$. It contains $S O(3)$ as a subgroup and has fundamental spinor representation that changes sign under a full rotation. The vector $\chi^{\dagger} \boldsymbol{\sigma} \chi$ provides a spin-1 representation of $S U(2)$ which is equivalent to the fundamental representation of $S O(3)$. All higher representations of the extended rotation group, i.e. describing higher spins, can be found as higher products of spinor representations.

### 5.8 Orbital angular momentum and spherical harmonics

A little thought shows that it is only rotations of systems with half-integer spins which pick up an extra minus sign after a full rotation. Integer-spin systems don't do that. For these systems we can also calculate the representation matrices using the same procedure as for spin- $1 / 2$. For instance for spin- 1 , the representation space is three-dimensional and we find for the three angular momentum generators

$$
L_{x}=\hbar \sqrt{\frac{1}{2}}\left(\begin{array}{ccc}
0 & 1 & 0  \tag{5.90}\\
1 & 0 & 1 \\
0 & 1 & 0
\end{array}\right) \quad L_{y}=\hbar \sqrt{\frac{1}{2}}\left(\begin{array}{rrr}
0 & -i & 0 \\
i & 0 & -i \\
0 & i & 0
\end{array}\right) \quad L_{z}=\hbar\left(\begin{array}{rrr}
1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & -1
\end{array}\right)
$$

The matrix $L_{z}$ follows directly from the eigenvalues in (5.67) while $L_{x}$ and $L_{y}$ follow from using (5.70) in the same way as we did for spin- $1 / 2$. We notice the useful relation $\left(L_{z} / \hbar\right)^{3}=L_{z} / \hbar$ which is also seen to be satisfied for $L_{x}$ and $L_{y}$. Again it should be stressed that it is only valid here for a spin-1 system.

It is tempting to relate this three-dimensional representation space to the one provided by the three components of the position vector $\mathbf{r}=x \mathbf{e}_{x}+y \mathbf{e}_{y}+z \mathbf{e}_{z}$ of a particle. We have already seen that in the coordinate representation the rotation generators are given by the differential operators $L_{x}=-i \hbar\left(y \partial_{z}-z \partial_{y}\right), L_{y}=-i \hbar\left(z \partial_{x}-x \partial_{z}\right)$ and $L_{z}=-i \hbar\left(x \partial_{y}-y \partial_{x}\right)$ where $\partial_{i} \equiv \partial / \partial x_{i}$. Using instead spherical coordinates with $x=r \sin \theta \cos \phi, y=r \sin \theta \sin \phi$ and $z=r \cos \theta$ we can express the partial derivatives $\partial_{x}, \partial_{y}$ and $\partial_{z}$ in terms of $\partial_{r}, \partial_{\theta}$ and $\partial_{\phi}$ after some work. It then follows that $L_{z}=-i \hbar \partial_{\phi}$ while the ladder operators become

$$
\begin{equation*}
L_{ \pm}=\hbar e^{ \pm i \phi}\left( \pm \frac{\partial}{\partial \theta}+i \cot \theta \frac{\partial}{\partial \phi}\right) \tag{5.91}
\end{equation*}
$$

From (5.63) we then also find for the total angular momentum operator

$$
\begin{equation*}
\mathbf{L}^{2}=-\hbar^{2}\left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial}{\partial \theta}\right)+\frac{1}{\sin ^{2} \theta} \frac{\partial^{2}}{\partial \phi^{2}}\right] \tag{5.92}
\end{equation*}
$$

All dependence on $\partial_{r}$ goes away since $r$ remains constant under rotations. The angular momentum eigenfunction $\langle\mathbf{r} \mid l, m\rangle$ are therefore functions of only the two angles $\theta$ and $\phi$. They are the spherical harmonics

$$
\begin{equation*}
Y_{l m}(\theta, \phi)=\langle\theta, \phi \mid l, m\rangle \tag{5.93}
\end{equation*}
$$

we shall now derive. The eigenvalue equation (5.67) now becomes the simple differential equation $-i \hbar \partial_{\phi} Y_{l m}=\hbar m Y_{l m}$ with the solution

$$
\begin{equation*}
Y_{l m}(\theta, \phi)=e^{i m \phi} y_{l}(\theta) \tag{5.94}
\end{equation*}
$$

This remaining function $y_{l}(\theta)$ will folow from the highest state which as defined by $\widehat{L}_{+}|l, l\rangle=0$. The corresponding wave function $Y_{l l}$ therefore satisfies the differential equation $\left(\partial_{\theta}+i \cot \theta \partial_{\phi}\right) Y_{l l}(\theta, \phi)=0$ or

$$
\begin{equation*}
\left(\partial_{\theta}-l \cot \theta\right) y_{l}(\theta)=0 \tag{5.95}
\end{equation*}
$$

using the raising operator in (5.91) and the split (5.94). Rewriting it as $d y_{l} / y_{l}=$ $l \cot \theta d \theta$, we can easily integrate both sides to give $\ln y_{l}=l \ln \sin \theta+\ln N_{l}$ where $N_{l}$ is an integration constant. We thus have the result $y_{l}(\theta)=N_{l} \sin ^{l} \theta$ where $N_{l}$ now appears as a normalization constant determined by

$$
\begin{equation*}
\left\langle l, m \mid l^{\prime}, m^{\prime}\right\rangle=\int d \Omega Y_{l m}^{*} Y_{l^{\prime} m^{\prime}}=\delta_{l l^{\prime}} \delta_{m m^{\prime}} \tag{5.96}
\end{equation*}
$$

where $d \Omega=\sin \theta d \theta d \phi$ is the solid angle differential. For our solution we therefore have

$$
\left|N_{l}\right|^{2} \int_{0}^{2 \pi} d \phi \int_{0}^{\pi} d \theta \sin \theta \sin ^{2 l} \theta=1
$$

where we can use the tabulated integral

$$
\begin{equation*}
\int_{0}^{\pi} d \theta \sin ^{p} \theta=\sqrt{\pi} \frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{p+2}{2}\right)} \tag{5.97}
\end{equation*}
$$

When $p=2 l+1$ is a positive integer, this can be simplified to

$$
\begin{equation*}
\int_{0}^{\pi} d \theta \sin ^{2 l+1} \theta=\frac{2^{l+1} l!}{(2 l+1)!!}=\frac{2\left(2^{l} l!\right)^{2}}{(2 l+1)!} \tag{5.98}
\end{equation*}
$$

since $\Gamma(1 / 2)=\sqrt{\pi}$. The normalization constant is thus found to be

$$
\begin{equation*}
N_{l}=\frac{(-1)^{l}}{2^{l} l!} \sqrt{\frac{(2 l+1)!}{4 \pi}} \tag{5.99}
\end{equation*}
$$

where the phase factor in front is a standard convention.
Having found the highest angular momentum eigenfunction, the others will now follow from using the lowering operator $L_{-}$together with the master formula (5.70). It gives

$$
L_{-} Y_{l l}=\hbar \sqrt{2 l \cdot 1} Y_{l, l-1}
$$

from which $Y_{l, l-1}$ is obtained. One more step down gives

$$
L_{-}^{2} Y_{l l}=\hbar^{2} \sqrt{2 l \cdot(2 l-1)} \sqrt{1 \cdot 2} Y_{l, l-2}
$$

and in general

$$
L_{-}^{l-m} Y_{l l}=\hbar^{l-m} \sqrt{2 l \cdot(2 l-1) \cdots(l+m+1)} \sqrt{1 \cdot 2 \cdot 3 \cdots(l-m)} Y_{l m}
$$

Thus one has

$$
\begin{equation*}
Y_{l m}(\theta, \phi)=\sqrt{\frac{(l+m)!}{(2 l)!(l-m)!}}\left(\frac{L_{-}}{\hbar}\right)^{l-m} Y_{l l}(\theta, \phi) \tag{5.100}
\end{equation*}
$$

where the differential operator $L_{-}$is given by (5.91) and $Y_{l l}$ as above.
One can simplify somewhat this result by observing that

$$
\begin{aligned}
\left(\frac{L_{-}}{\hbar}\right) y_{l} e^{i m \phi} & =e^{-i \phi}\left(-\frac{\partial}{\partial \theta}+i \cot \theta \frac{\partial}{\partial \phi}\right) y_{l} e^{i m \phi} \\
& =-e^{i(m-1) \phi}\left(y_{l}^{\prime}+m y_{l} \cot \theta\right)
\end{aligned}
$$

is quite similar to

$$
\left(\frac{d}{d \cos \theta}\right) y_{l} \sin ^{m} \theta=-\left(y_{l}^{\prime}+m y_{l} \cot \theta\right) \sin ^{m-1} \theta
$$

where the dash derivative is with respect to $\theta$. Repeating now this operation $l-m$ times, one thus obtains

$$
\begin{equation*}
\left(\frac{L_{-}}{\hbar}\right)^{l-m} e^{i m \phi} \sin ^{l} \theta=e^{i m \phi} \sin ^{-m} \theta\left(\frac{d}{d \cos \theta}\right)^{l-m} \sin ^{2 l} \theta \tag{5.101}
\end{equation*}
$$

which is what we need in (5.94). Including the normalization constant (5.99) in $Y_{l l}$, we get the final result for the spherical harmonic functions

$$
\begin{equation*}
Y_{l m}(\theta, \phi)=\frac{(-1)^{l}}{2^{l} l!} \sqrt{\frac{2 l+1}{4 \pi} \cdot \frac{(l+m)!}{(l-m)!}} e^{i m \phi} \sin ^{-m} \theta\left(\frac{d}{d \cos \theta}\right)^{l-m} \sin ^{2 l} \theta \tag{5.102}
\end{equation*}
$$

From this construction it can be shown that $Y_{l,-m}(\theta, \phi)=(-1)^{m} Y_{l m}^{*}(\theta, \phi)$ which is a useful symmetry relation.
The $l=0$ spherical harmonic is just the constant $Y_{00}=1 / \sqrt{4 \pi}$. When $l=1$ there are three functions

$$
\begin{align*}
Y_{1,0} & =\sqrt{\frac{3}{4 \pi}} \cos \theta=\sqrt{\frac{3}{4 \pi}} \frac{z}{r}  \tag{5.103}\\
Y_{1, \pm 1} & =\mp \sqrt{\frac{3}{8 \pi}} e^{ \pm i \phi} \sin \theta=\mp \sqrt{\frac{3}{8 \pi}} \frac{x \pm i y}{r} \tag{5.104}
\end{align*}
$$

We thus see that the three coordinates $(x, y, z)$ provide a three-dimensional representation space for angular momentum $l=1$, just as advertised in the beginning of this section.

When the azimuthal quantum number $m=0$ the spherical harmonics are seen to be polynomials in $\cos \theta$. More accurately, writing

$$
\begin{equation*}
Y_{l 0}(\theta, \phi)=\sqrt{\frac{2 l+1}{4 \pi}} P_{l}(\cos \theta), \tag{5.105}
\end{equation*}
$$

the Legendre polynomials are defined by

$$
\begin{equation*}
P_{l}(x)=\frac{1}{2^{l} l!}\left(-\frac{d}{d x}\right)^{l}\left(1-x^{2}\right)^{l} \tag{5.106}
\end{equation*}
$$

when comparing with (5.102). Their normalization follows from (5.96) which now gives

$$
\begin{equation*}
\int_{-1}^{1} d x P_{l}(x) P_{l^{\prime}}(x)=\frac{2}{2 l+1} \delta_{l l^{\prime}} \tag{5.107}
\end{equation*}
$$

It is straightforward to calculate the lowest of these functions from (5.106). One finds $P_{0}(x)=1$ while

$$
\begin{aligned}
& P_{1}(x)=-\frac{1}{2} \frac{d}{d x}\left(1-x^{2}\right)=x \\
& P_{2}(x)=\frac{1}{8} \frac{d^{2}}{d x^{2}}\left(1-x^{2}\right)^{2}=\frac{1}{8} \frac{d}{d x}\left(4 x^{3}-4 x\right)=\frac{1}{2}\left(3 x^{2}-1\right)
\end{aligned}
$$

and so on. We see from here that $P_{l}(1)=1$ which is true also for all higher values of $l$. These functions are even in $x$ when $l$ is even and odd when $l$ is odd. This can be summed up in the relation

$$
\begin{equation*}
P_{l}(-x)=(-1)^{l} P_{l}(x) \tag{5.108}
\end{equation*}
$$

which we will return to in the following section.

### 5.9 Spherical symmetric potentials

We now have the tools to investigate the quantum mechanics of a particle with mass $\mu$ in a spherical symmetric potential $V=V(r)$. Using spherical coordinates, the Hamiltonian

$$
\begin{equation*}
H=-\frac{\hbar^{2}}{2 \mu} \nabla^{2}+V(r) \tag{5.109}
\end{equation*}
$$

will act on wave function $\psi(\mathbf{r})=\psi(r, \theta, \phi)$. We thus need

$$
\begin{equation*}
\nabla^{2}=\frac{\partial^{2}}{\partial r^{2}}+\frac{2}{r} \frac{\partial}{\partial r}+\frac{1}{r^{2}}\left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial}{\partial \theta}\right)+\frac{1}{\sin ^{2} \theta} \frac{\partial^{2}}{\partial \phi^{2}}\right] \tag{5.110}
\end{equation*}
$$

The last, angular part is seen to be just the total angular momentum operator $\mathbf{L}^{2}$ in (5.92). This can be understood from the composition of the classical momentum $\mathbf{p}$ into a radial part and an angular part which follows from $(\mathbf{r} \wedge \mathbf{p})^{2}=r^{2} p^{2}-(\mathbf{r} \cdot \mathbf{p})^{2}$ so that

$$
\begin{equation*}
\mathbf{p}^{2}=p_{r}^{2}+\frac{1}{r^{2}} \mathbf{L}^{2} \tag{5.111}
\end{equation*}
$$

where the radial part $p_{r}=(\mathbf{r} \cdot \mathbf{p}) / r$. In the quantum theory $\mathbf{p}$ becomes the differential operator $\mathbf{p}=-i \hbar \boldsymbol{\nabla}$ which no longer commutes with the radius $r$. In order for the operator $p_{r}$ to be Hermitean, it must therefore be defined as

$$
\begin{equation*}
p_{r}=\frac{1}{2}\left(\frac{1}{r} \mathbf{r} \cdot \mathbf{p}+\mathbf{p} \cdot \mathbf{r} \frac{1}{r}\right)=-i \hbar \frac{1}{r} \frac{\partial}{\partial r} r \tag{5.112}
\end{equation*}
$$

A similar, short calculation then shows that

$$
p_{r}^{2}=-\frac{\hbar^{2}}{r^{2}} \frac{\partial}{\partial r}\left(r^{2} \frac{\partial}{\partial r}\right)=-\hbar^{2}\left(\frac{\partial^{2}}{\partial r^{2}}+\frac{2}{r} \frac{\partial}{\partial r}\right)
$$

which is the first part of the Laplacian (5.110). Its structure therefore corresponds to the classical decomposition (5.111).

The stationary states of the particle in this potential will now be given by the solutions of the differential equation

$$
\begin{equation*}
\left[-\frac{\hbar^{2}}{2 \mu}\left(\frac{\partial^{2}}{\partial r^{2}}+\frac{2}{r} \frac{\partial}{\partial r}\right)+\frac{\mathbf{L}^{2}}{2 \mu r^{2}}+V(r)\right] \psi(\mathbf{r})=E \psi(\mathbf{r}) \tag{5.113}
\end{equation*}
$$

It can be solved along the same lines as the cylindrical symmetric systems considered earlier in this chapter. Since the Hamiltonian is rotational invariant so that $[H, \mathbf{L}]=$ 0 , the energy eigenfunctions $\psi(\mathbf{r})$ are also eigenfunctions of the angular momentum operator $\mathbf{L}$. We can therefore write $\psi(\mathbf{r})=R(r) Y_{l m}(\theta, \phi)$ where the radial function $R(r)$ is seen to satisfy

$$
\begin{equation*}
\left[-\frac{\hbar^{2}}{2 \mu}\left(\frac{d^{2}}{d r^{2}}+\frac{2}{r} \frac{d}{d r}\right)+\frac{l(l+1) \hbar^{2}}{2 \mu r^{2}}+V(r)\right] R(r)=E R(r) \tag{5.114}
\end{equation*}
$$

Now since

$$
\frac{d^{2}}{d r^{2}}(r R)=r \frac{d^{2} R}{d r^{2}}+2 \frac{d R}{d r}
$$

the equation simplifies by introducing the new function $u(r)=r R(r)$ satisfying the simpler equation

$$
\begin{equation*}
-\frac{\hbar^{2}}{2 \mu} \frac{d^{2} u}{d r^{2}}+\left[V(r)+\frac{l(l+1) \hbar^{2}}{2 \mu r^{2}}\right] u=E u \tag{5.115}
\end{equation*}
$$

It has exactly the same form as the one-dimensional motion of a particle in the effective potential

$$
\begin{equation*}
V_{e f f}(r)=V(r)+\frac{l(l+1) \hbar^{2}}{2 \mu r^{2}} \tag{5.116}
\end{equation*}
$$

The last term here is the contribution from the centrifugal motion. If the potential $V(r)$ is less singular than $1 / r^{2}$, this centrifugal part will therfore determine the wave function as the radius $r \rightarrow 0$. In fact, if we neglect the $E$ and $V$ compared with the centrifugal term, the differential equation (5.115) simplifies to

$$
\frac{d^{2} u}{d r^{2}}=\frac{1}{r^{2}} l(l+1) u
$$

This has solutions of the form $u \propto r^{s}$ which is verified by substution. This leads to the condition $s(s-1)=l(l+1)$ with the two solutions $s=-l$ or $s=l+1$. But the negative solution means that the wave function diverges in the limit $r \rightarrow 0$ which we want to avoid. We must therefore choose the other solution, giving $u \propto r^{l+1}$ in this region. Separating out this part from the general solution, it must of the form

$$
\begin{equation*}
u(r) \propto r^{l+1} v(r) \tag{5.117}
\end{equation*}
$$

where now $v(r)$ remains to be determined.
Only for a few special potentials $V(r)$ can the differential equation (5.115) be solved analytically. The simplest case is a free particle for which $V=0$. Writing then the energy as $E=\hbar^{2} k^{2} / 2 \mu$ so that it is expressed by the unknown quantity $k$, the radial equation (5.115) becomes

$$
\begin{equation*}
\frac{d^{2} u}{d x^{2}}-\frac{l(l+1)}{x^{2}} u+u=0 \tag{5.118}
\end{equation*}
$$

where now $x=k r$ is a dimensionless radial variable. For $l=0$ this is now just an ordinary, harmonic equation with the solutions $u=\sin x$ or $u=\cos x$. Again we reject this last possibilty since it gives a radiel function $R=u / r$ diverging when $r \rightarrow 0$. The regular solution is therefore

$$
\begin{equation*}
R_{0}(r)=\frac{\sin k r}{k r} \tag{5.119}
\end{equation*}
$$

when we ignore any normalization constants.
Solutions with higher angular momentum $l>0$ can be derived by rewriting (5.118) as a differential equation for the function $v=u / x^{l+1}$. One then finds the equivalent equation

$$
\begin{equation*}
v^{\prime \prime}+\frac{2}{x}(l+1) v+v=0 \tag{5.120}
\end{equation*}
$$

This now gives a recursion relation when we introduce the new function $w=-v^{\prime} / x$ and taking the derivative, i.e.

$$
(x w)^{\prime \prime}+\frac{2}{x}(l+1)(x w)^{\prime}+\left[1-\frac{2}{x^{2}}(l+1)\right] x w=0
$$

Simplifying, one then obtains

$$
w^{\prime \prime}+\frac{2}{x}(l+2) w+w=0
$$

which is just the same as (5.120), but now for $l+1$. In this way we have established the recursion relation

$$
\begin{equation*}
v_{l+1}=-\frac{1}{x} \frac{d v_{l}}{d x} \tag{5.121}
\end{equation*}
$$

with $v_{0}=\sin x / x$ as calculated above. It gives immediately

$$
v_{l}=\left(-\frac{1}{x} \frac{d}{d x}\right)^{l} \frac{\sin x}{x}
$$

and from (5.117) therefore for the full radial function

$$
\begin{equation*}
R_{l}(x)=x^{l}\left(-\frac{1}{x} \frac{d}{d x}\right)^{l} \frac{\sin x}{x} \tag{5.122}
\end{equation*}
$$

These functions are called spherical Bessel functions and denoted by $j_{l}(x)$. For the first ones we have

$$
\begin{aligned}
& j_{0}(x)=\frac{\sin x}{x} \\
& j_{1}(x)=\frac{\sin x}{x^{2}}-\frac{\cos x}{x} \\
& j_{2}(x)=\left(\frac{3}{x^{3}}-\frac{1}{x}\right) \sin x-\frac{3}{x^{2}} \cos x
\end{aligned}
$$

and so on. That they are solutions to the Bessel equation (5.21) is not so surprising. When we write the radial equation (5.114) with $V=0$ and using the dimensionless radius $x=k r$, it becomes $x^{2} R^{\prime \prime}+2 x R^{\prime}+\left(x^{2}-l(l+1)\right) R=0$ which show that they are quite similar. In fact, they are related by

$$
\begin{equation*}
j_{l}(x)=\sqrt{\frac{\pi}{2 x}} J_{l+1 / 2}(x) \tag{5.123}
\end{equation*}
$$

For this reason the spherical Bessel function are sometimes also called half-integer Bessel functions.

In spherical coordinates we now have found the general form

$$
\psi_{\mathbf{k}}(\mathbf{r})=\sum_{l, m} c_{l, m} j_{l}(k r) Y_{l m}(\theta, \phi)
$$

for the wave function of a free particle with energy $E=(\hbar k)^{2} / 2 \mu$. The expansion coefficients are determined from the requirement that this infinite sum must describe
the plane wave $\psi_{\mathbf{k}}=e^{i \mathbf{k} \cdot \mathbf{r}}$ in Cartesian coordinates. They are simplest to calculate when the wave number $\mathbf{k}$ is along the $z$-axis. The right-hand side can then not depend on the azimuthal angle $\phi$ so that the quantum number $m=0$. From (5.105) we can then express the spherical harmonic $Y_{l 0}$ by the Legendre polynomial $P_{l}$ giving

$$
e^{i k z}=\sum_{l=0}^{\infty} C_{l} j_{l}(k r) P_{l}(\cos \theta)
$$

with $z=r \cos \theta$. Using the orthogonality relation (5.107) between Legendre polynomials and the behaviour of $j_{l}(k r)$ in the limit $r \rightarrow 0$, one can calculate the expansion coefficients $C_{l}$. In this way we end up with the plane-wave expansion

$$
\begin{equation*}
e^{i k r \cos \theta}=\sum_{l=0}^{\infty} i^{l}(2 l+1) j_{l}(k r) P_{l}(\cos \theta) \tag{5.124}
\end{equation*}
$$

It has many useful applications as in the theory of particle scattering.

### 5.10 The spherical well

A model for protons and neutrons confined within the nucleus, is to assume that they move freely within the spherical well potential

$$
V(r)= \begin{cases}0, & r<a  \tag{5.125}\\ \infty, & r>a\end{cases}
$$

More realistically one should assume a finite depth, but that problem is bit more difficult to investigate. Inside the well where the potential $V=0$, a confined particle will be described by the the radial wavefunctions $R(r)=j_{l}(k r)$. When the potential becomes infinite at the wall of the well at radius $r=a$, the wave function must be zero resulting in the condition $j_{l}(k a)=0$. The wave number $k$ for the mode with angular momentum $l$ is therefore quantized with the possible values $k=x_{l n} / a$ where $x_{l n}$ is the $n$ 'th root of the Bessel function $j_{l}(x)$. As a result we immediately have the energy spectrum given as

$$
\begin{equation*}
E_{n l}=\frac{\hbar^{2}}{2 \mu a^{2}} x_{l n}^{2} \tag{5.126}
\end{equation*}
$$

It is customary to call $n_{r}=n-1$ for the radial quantum number, giving the number of zeros in the radial wave function. The $s$-states have $l=0$ and therefore the wave function (5.119) with zeros for $x_{0 n}=n \pi$, i.e. $x_{01}=3.14, x_{02}=6.28$ and so on. For the $p$-states with $l=1$ the lowest zeros are at $x_{11}=4.49$ and $x_{12}=7.73$, the lowest $l=2$ or $d$-state is given by $x_{21}=5.76$ while the first $f$-level with $l=3$ has $x_{31}=6.99$. These energy levels are shown in Fig. 5.4 with labels $(n, l)$.

This spherical well potential is quite similar to the three-dimensional harmonic oscillator potential $V=\mu \omega^{2} r^{2} / 2$ which also gives an analytically solvable problem. One then proceeds along the same lines as for the planar oscillator in the beginning of this section. The eigenfunctions will have a form similar to the two-dimensional solution (5.32), but now in terms of an associated Laguerre polynomial of half-integer order. From the


Figure 5.4: Lowest energy levels in a spherical well in units of $E_{0}=\hbar^{2} / 2 \mu a^{2}$.
formula for the energy levels $E_{n l}=\left(2 n_{r}+l+3 / 2\right) \hbar \omega$ where the quantum number $n_{r}=n-1$ again gives the number of zeros in the radial wave function, we see that the $2 s$ level now is degenerate with $1 d$, while $2 p$ is degenerate with $1 f$. Needless to say, the resulting energy level diagram is the same as obtained previously in Cartesian coordinates.

Central potentials are also invariant under the discrete symmetry $\mathbf{r} \rightarrow-\mathbf{r}$ which is now the parity operation. This is also an invariance of the full Hamiltonian and we can thus assign the parity eigenvalue $P= \pm 1$ to each energy eigenstate. Under this transformation the spherical angles change as $\theta \rightarrow \pi-\theta$ and $\phi \rightarrow \pi+\phi$. From the definition (5.102) of the spherical harmonics, we therefore see that they change as $Y_{l m} \rightarrow$ $(-1)^{m}(-1)^{l-m} Y_{l m}=(-1)^{l} Y_{l m}$. The radial wave function is unchanged. We thus find that the parity quantum number of a state with angular momentum quantum number $l$ is simply $(-1)^{l}$.

### 5.11 Hydrogen atoms

In a hydrogen-like atom we have one particle with charge $e$ and mass $m_{1}$ interacting with the nuclues with mass $m_{2}$ via the Coulomb potential. If $Z$ is the number of positive charges on the nucleus, the full Hamiltonian of the atom is then

$$
\begin{equation*}
H_{t o t}=\frac{\mathbf{p}_{1}^{2}}{2 m_{1}}+\frac{\mathbf{p}_{2}^{2}}{2 m_{2}}-\frac{Z e^{2}}{4 \pi \epsilon_{0}\left|\mathbf{r}_{1}-\mathbf{r}_{2}\right|} \tag{5.127}
\end{equation*}
$$

As in classical mechanics, this is separated into the motion of the center-of-mass with coordinates

$$
\begin{equation*}
\mathbf{R}=\frac{m_{1} \mathbf{r}_{1}+m_{2} \mathbf{r}_{2}}{m_{1}+m_{2}} \tag{5.128}
\end{equation*}
$$

and momentum $\mathbf{P}=\mathbf{p}_{1}+\mathbf{p}_{2}$ plus a relative motion described by the coordinate difference $\mathbf{r}=\mathbf{r}_{1}-\mathbf{r}_{2}$. The Hamiltonian then becomes $H_{t o t}=H_{C M}+H$ where $H_{C M}=\mathbf{P}^{2} / 2 M$ governs the motion of the center of mass. It will therefore move as a free particle with mass $M=m_{1}+m_{2}$. Of more interest is the relative motion described by

$$
\begin{equation*}
H=\frac{1}{2 \mu} \mathbf{p}^{2}-\frac{Z e^{2}}{4 \pi \epsilon_{0} r} \tag{5.129}
\end{equation*}
$$

where $\mu$ is the reduced mass

$$
\begin{equation*}
\mu=\frac{m_{1} m_{2}}{m_{1}+m_{2}} \tag{5.130}
\end{equation*}
$$

In an ordinary hydrogen atom with an electron bound to a proton $m_{1}=m_{e}$ and $m_{2}=$ $m_{p} \gg m_{e}$ so that $\mu \approx m_{e}$, while in positronium $\mu=m_{e} / 2$.
The Hamiltonian (5.129) is spherical symmetric and can be solved within the framework already established. The reduced radial equatiom (5.115) now becomes

$$
\begin{equation*}
\frac{\hbar^{2}}{2 \mu} \frac{d^{2} u}{d r^{2}}+\left[E+\frac{Z e^{2}}{4 \pi \epsilon_{0} r}-\frac{l(l+1) \hbar^{2}}{2 \mu r^{2}}\right] u=0 \tag{5.131}
\end{equation*}
$$

From (5.117) we already know know that a physical meaningful solution of this differential equation for small separations $r \rightarrow 0$ must behave as $u(r) \propto r^{l+1}$ since the centrifugal term then dominates in the effective potential. Since the Coulomb potential is negative, and goes to zero at large separations, any bound state must have negative energy. In this region it simplifies to

$$
\frac{d^{2} u}{d r^{2}}+\frac{2 \mu E}{\hbar^{2}} u=0
$$

The bound state has a wave function which must go to zero when $r \rightarrow \infty$. We thus find the acceptable solution $u(r) \propto e^{-\kappa r}$ in this region when introducing $\kappa^{2}=-2 \mu E / \hbar^{2}$.

In order to simplify the differential equation (5.131), we now first introduce a dimensionless radial variable $\rho=2 \kappa r$. It then becomes

$$
\begin{equation*}
\frac{d^{2} u}{d \rho^{2}}+\left[\frac{\lambda}{\rho}-\frac{l(l+1)}{\rho^{2}}-\frac{1}{4}\right] u=0 \tag{5.132}
\end{equation*}
$$

where

$$
\begin{equation*}
\lambda=\frac{Z e^{2}}{4 \pi \epsilon_{0} \hbar} \sqrt{\frac{\mu}{-2 E}} \tag{5.133}
\end{equation*}
$$

Next, we separate out the established behaviour of the wave function in the limits $\rho \rightarrow 0$ and $\rho \rightarrow \infty$ by writing

$$
u(r)=\rho^{l+1} e^{-\rho / 2} L(\rho)
$$

We are then left with a differential equation for the function $L(\rho)$ which is found by substituting this expression into (5.131). The result is

$$
\begin{equation*}
\rho L^{\prime \prime}+(2 l+2-\rho) L^{\prime}+(\lambda-l-1) L=0 \tag{5.134}
\end{equation*}
$$

Luckily, this is Kummer's equation (5.27) we already have met in connection with the quantization of the planar oscillator. It has solutions which are associated Laguerre polynomials bounded in the limit $\rho \rightarrow \infty$ only when the coefficient $\lambda-l-1 \equiv n_{r}$ is a positive integer or zero, i.e. $n_{r}=0,1,2,3, \cdots$. This is the radial quantum number, giving the number of zeros in the corresponding polynomial $L(\rho)$.
The quantity (5.133) is now an integer, $\lambda=l+1+n_{r} \equiv n$ called the principal quantum number, with values $n=l+1, l+2, l+3, \cdots$. It results in the quantized energy levels

$$
\begin{equation*}
E_{n}=-\frac{1}{2} \mu\left(\frac{Z e^{2}}{4 \pi \epsilon_{0} \hbar}\right)^{2} \frac{1}{n^{2}} \tag{5.135}
\end{equation*}
$$

which is the result found by Bohr ten years before modern quantum mechanics. Denoting each each eigenstate by the quantum numbers $(n l)$, the lowest levels are shown in Fig.5.5. All states with the same principal quantum number $n$ and angular momentum quantum number in the interval $0 \leq l \leq n-1$ have the same energy. Since each eigenstate with


Figure 5.5: Lowest energy levels in the hydrogen atom.
a definite value of $l$, has a degeneracy of $2 l+1$, the total degeneracy of an energy level $E_{n}$ is

$$
\begin{equation*}
\sum_{l=0}^{n-1}(2 l+1)=2 \sum_{l=0}^{n-1} l+n=n(n-1)+n=n^{2} \tag{5.136}
\end{equation*}
$$

This accidental degeneracy is due to some extra symmetry in the Coulomb potential on top of the three-dimensional, rotational invariance. In fact, the full symmetry corresponds to rotations in an abstract, four-dimensional space.

A more compact and useful version of the Bohr formula (5.135) is

$$
\begin{equation*}
E_{n}=-\frac{1}{2} \mu c^{2} \alpha^{2} \frac{Z^{2}}{n^{2}} \tag{5.137}
\end{equation*}
$$

when we introduce the velocity of light $c$ and the fine structure constant

$$
\begin{equation*}
\alpha=\frac{e^{2}}{4 \pi \epsilon_{0} \hbar c}=\frac{1}{137.036 \ldots} \tag{5.138}
\end{equation*}
$$

In a later chapter we will understand better why this name is appropriate. The corresponding eigenfunctions are

$$
\begin{equation*}
\psi_{n l m}(\mathbf{r})=N_{n l} \rho^{l} e^{-\rho / 2} L_{n-l-1}^{2 l+1}(\rho) Y_{l m}(\theta, \phi) \tag{5.139}
\end{equation*}
$$

where $N_{n l}$ is a normalization constant. Notice that the lower index on the Laguerre polynomial is the radial quantum number.
We are here most interested in the case of ordinary atoms where the bound particle is an electron with mass $m_{e}$. Taking the nuclear mass to infinity, we will then have a reduced mass $\mu=m_{e}$. The dimensionless coordinate $\rho=2 \kappa r$ then becomes

$$
\begin{equation*}
\rho=2 r \sqrt{\frac{-2 m_{e} E}{\hbar^{2}}}=\frac{2 r}{n a} \tag{5.140}
\end{equation*}
$$

where $a=a_{0} / Z$ is called the radius of the atom in its ground state. It equals the Bohr radius $a_{0}=\hbar / \alpha m_{e} c=0.53 \times 10^{-10} \mathrm{~m}$ in the ordinary hydrogen atom with $Z=1$ for a proton. Notice that the larger $Z$ is, the smaller the atom becomes. The electron is then pulled more in towards the nucleus.
For hydrogen the ground state energy is $E_{1}=-m_{e} c^{2} \alpha^{2} / 2=-e^{2} / 8 \pi \epsilon_{0} a_{0}=-13.60 \mathrm{eV}$ from the value $m_{e}=0.511 \mathrm{MeV} / c^{2}$ of the electron mass. This sets the scale for all atomic energies. In this state the kinetic energy of the electron is $p^{2} / 2 m_{e}=e^{2} / 8 \pi \epsilon_{0} a_{0}$ and it therefore has a velocity $v / c=p / m_{e} c=\alpha$. With the above value for the fine structure constant, this is an enormous velocity, slightly less than one percent of the velocity of light. Relativistic corrections are therefore expected to come in at a magnitude $(v / c)^{2}=\alpha^{2} \approx 10^{-5}$ smaller than the ground state energy. Such corrections will be considered in more detail in a later chapter.
Writing now the eigenfunction in (5.139) as

$$
\begin{equation*}
\psi_{n l m}(r, \theta, \phi)=R_{n l}(r) Y_{l m}(\theta, \phi) \tag{5.141}
\end{equation*}
$$

with the radial function

$$
\begin{equation*}
R_{n l}(r)=N_{n l}(2 r / n a)^{l} e^{-r / n a} L_{n-l-1}^{2 l+1}(2 r / n a), \tag{5.142}
\end{equation*}
$$

the constant is determined from the orthonormality condition

$$
\begin{equation*}
\int d^{3} r \psi_{n l m}^{*}(\mathbf{r}) \psi_{n^{\prime} l^{\prime} m^{\prime}}(\mathbf{r})=\delta_{n n^{\prime}} \delta_{l l^{\prime}} \delta_{m m^{\prime}} \tag{5.143}
\end{equation*}
$$

Using the normalization (5.96) of the spherical harmonics, this simplifies to

$$
\begin{equation*}
\int_{0}^{\infty} d r r^{2} R_{n l}^{2}(r)=1 \tag{5.144}
\end{equation*}
$$

with the result

$$
\begin{equation*}
N_{n l}=\frac{2}{n^{2}} \sqrt{\frac{(n-l-1)!}{a^{3}(n+l)!}} \tag{5.145}
\end{equation*}
$$

for the normalization constant. In this way we find for the lowest radial functions

$$
\begin{align*}
R_{1 s} & =\sqrt{\frac{4}{a^{3}}} e^{-r / a}  \tag{5.146}\\
R_{2 s} & =\sqrt{\frac{1}{8 a^{3}}}\left(2-\frac{r}{a}\right) e^{-r / 2 a}  \tag{5.147}\\
R_{2 p} & =\sqrt{\frac{1}{24 a^{3}}} \frac{r}{a} e^{-r / 2 a} \tag{5.148}
\end{align*}
$$

using definition (5.28) of the associated Laguerre polynomials.
Later we will need the wave function at the origin $\mathbf{r}=0$. From (5.142) we see that it is only the $s$-states with $l=0$ which are non-zero at this point. It involves the value of the Laguerre polynomial $L_{n-1}^{1}(0)=n$ which follows from the definition (5.30). As a result we obtain

$$
\begin{equation*}
\psi_{n 00}(\mathbf{r}=0)=\frac{2}{n^{2}} \sqrt{\frac{(n-1)!}{a^{3} n!}} \frac{n}{\sqrt{4 \pi}}=\sqrt{\frac{1}{\pi a^{3} n^{3}}} \tag{5.149}
\end{equation*}
$$

### 5.12 Rotation matrices and Euler angles

So far we have only considered rotations being performed on a particle with coordinates $(x, y, z)$. in a fixed coordinate system. These are called active transformations. If the particle is initially in the state $|\psi\rangle$, it will then after the rotation $R(\boldsymbol{\alpha})$ be in the state $\left|\psi^{\prime}\right\rangle$ with the corresponding wavefunction $\psi^{\prime}(\mathbf{r})=\left\langle\mathbf{r} \mid \psi^{\prime}\right\rangle=\langle\mathbf{r}| \widehat{R}(\boldsymbol{\alpha})|\psi\rangle$. But now we can consider this expression in the passive sense where the system $|\psi\rangle$ remains fixed, but the coordinate system is rotated in the opposite way so that $\mathbf{r} \rightarrow R^{-1}(\boldsymbol{\alpha}) \mathbf{r}$. The wavefunction $\psi^{\prime}(\mathbf{r})$ represents then the same state, but in this new coordinate system.

If the quantum system under consideration is a single particle in the angular-momentum eigenstate $|l, m\rangle$, it will before the passive rotation be described by the wavefunction $Y_{l m}(\theta, \phi)=\langle\theta, \phi \mid l, m\rangle$ and afterwards by $Y_{l m}\left(\theta^{\prime}, \phi^{\prime}\right)=\left\langle\theta^{\prime}, \phi^{\prime} \mid l, m\right\rangle$ since it is found with the spherical angles $\left(\theta^{\prime}, \phi^{\prime}\right)$ in the new coordinate system. Inserting a complete set of angular-momentum eigenstates in this last matrix element, we then have

$$
\begin{align*}
Y_{l m}\left(\theta^{\prime}, \phi^{\prime}\right) & =\left\langle\mathbf{r}^{\prime} \mid l, m\right\rangle=\langle\mathbf{r}| \widehat{R}(\boldsymbol{\alpha})|l, m\rangle=\sum_{m^{\prime}=-l}^{l}\left\langle\mathbf{r} \mid l, m^{\prime}\right\rangle\left\langle l, m^{\prime}\right| \widehat{R}(\boldsymbol{\alpha})|l, m\rangle \\
& =\sum_{m^{\prime}=-l}^{l} Y_{l m^{\prime}}(\theta, \phi) D_{m^{\prime} m}^{(l)}(\boldsymbol{\alpha}) \tag{5.150}
\end{align*}
$$



Figure 5.6: Definition of the three Euler angles.
where the $(2 l+1) \times(2 l+1)$ matrix with elements

$$
\begin{equation*}
D_{m^{\prime} m}^{(l)}(\boldsymbol{\alpha})=\left\langle l, m^{\prime}\right| \widehat{R}(\boldsymbol{\alpha})|l, m\rangle=\left\langle l, m^{\prime}\right| e^{-i \boldsymbol{\alpha} \cdot \widehat{\mathbf{L}} / \hbar}|l, m\rangle \tag{5.151}
\end{equation*}
$$

is called the rotation matrix or Wigner function for spin $l$. Since the rotation operators $\widehat{R}(\boldsymbol{\alpha})$ are unitary, these matrices will also be unitary. Their physical content is perhaps more easily seen from the equivalent definition

$$
\begin{equation*}
\widehat{R}(\boldsymbol{\alpha})|l, m\rangle=\sum_{m^{\prime}=-l}^{l}\left|l, m^{\prime}\right\rangle D_{m^{\prime} m}^{(l)}(\boldsymbol{\alpha}) \tag{5.152}
\end{equation*}
$$

The ket vector on the left-hand side is an eigenvector of $\widehat{L}_{z}$ with the eigenvalue $m \hbar$ in the direction of the rotated $z$-axis. On the right-hand side this vector is decomposed onto basis eigenvectors for $\widehat{L}_{z}$ in the original $z$-direction. Thus $D_{m^{\prime} m}^{(l)}(\boldsymbol{\alpha})$ is the amplitude to measure the value $m^{\prime} \hbar$ along the $z$-axis when we know that the spin points in the direction of the rotated $z$-axis with component $m \hbar$. This observation enables many simple explanations of spin effects in different experimential situations in nuclear and elementary particle physics.

Calculation of the above rotation matrices is most easily done parametrisizing the rotations with the three Euler angles used in classical mechanics instead of the rotation vector $\boldsymbol{\alpha}=\alpha \mathbf{n}$ corresponding to a single rotation of $\alpha$ around the axis $\mathbf{n}$. A general rotation of the original axis system $(x, y, z)$ is then defined by the three Euler rotations

1. A counter-clockwise rotation of $\alpha$ around the $z$-axis so that the original axis system goes into the new system ( $x^{\prime \prime}, y^{\prime \prime}, z^{\prime \prime}=z$ ).
2. A counter-clockwise rotation of $\beta$ around the $y^{\prime \prime}$-axis. This carries the $z^{\prime \prime}$-axis into the $z^{\prime}$-axis.
3. A counter-clockwise rotation of $\gamma$ around the $z^{\prime}$-axis. This carries the $x^{\prime \prime}$-axis into the $x^{\prime}$-axis and $y^{\prime \prime}$-axis into the $y^{\prime}$-axis.

The final axis system $\left(x^{\prime}, y^{\prime}, z^{\prime}\right)$ is thus obtained by the combined rotations

$$
\begin{equation*}
\widehat{R}(\alpha, \beta, \gamma)=e^{-i \gamma \widehat{L}_{z^{\prime}} / \hbar} e^{-i \beta \widehat{L}_{y^{\prime \prime}} / \hbar} e^{-i \alpha \widehat{L}_{z} / \hbar} \tag{5.153}
\end{equation*}
$$

But now since the second rotation is around the rotated $y$-axis, we can make use of (5.50) to write it as a rotation about the original $y$-axis, i.e.

$$
e^{-i \beta \widehat{L}_{y^{\prime \prime}} / \hbar}=e^{-i \alpha \widehat{L}_{z} / \hbar} e^{-i \beta \widehat{L_{y}} / \hbar} e^{i \alpha \widehat{L}_{z} / \hbar}
$$

Similarly, $e^{-i \gamma \widehat{L}_{z^{\prime}} / \hbar}$ is the transform of $e^{-i \gamma \widehat{L_{z}} / \hbar}$ under the previous rotation that carried the $z$-axis into the $z^{\prime}$-axis, i.e.

$$
e^{-i \gamma \widehat{L_{z^{\prime}}} / \hbar}=e^{-i \beta \widehat{L}_{y^{\prime \prime}} / \hbar} e^{-i \gamma \widehat{L}_{z} / \hbar} e^{i \widehat{\beta \hat{L}_{y^{\prime \prime}} / \hbar}}
$$

Thus we find for the full rotation (5.153)

$$
\begin{align*}
\widehat{R}(\alpha, \beta, \gamma) & =e^{-i \beta \widehat{L}_{y^{\prime \prime}} / \hbar} e^{-i \gamma \widehat{L}_{z} / \hbar} e^{-i \alpha \widehat{L}_{z} / \hbar} \\
& =e^{-i \alpha \widehat{L}_{z} / \hbar} e^{-i \beta \widehat{L_{y}} / \hbar} e^{-i \gamma \widehat{L}_{z} / \hbar} \tag{5.154}
\end{align*}
$$

It corresponds to three separate rotations now performed in the original and fixed axis system $(x, y, z)$. This represents a big simplification.
In the Euler parametrization the inverse rotation defined by $\widehat{R} \widehat{R}^{-1}=\widehat{R}^{-1} \widehat{R}=1$ is now $\widehat{R}^{-1}(\alpha, \beta, \gamma)=\widehat{R}(-\gamma,-\beta,-\alpha)$ which is easily verified. These operators are elements of the three-dimensional rotation group $S O(3)$ where the product of two rotations is equivalent to a new rotation, i.e.

$$
\begin{equation*}
\widehat{R}(\alpha, \beta, \gamma) \widehat{R}\left(\alpha^{\prime}, \beta^{\prime}, \gamma^{\prime}\right)=\widehat{R}\left(\alpha^{\prime \prime}, \beta^{\prime \prime}, \gamma^{\prime \prime}\right) \tag{5.155}
\end{equation*}
$$

where the angles $\left(\alpha^{\prime \prime}, \beta^{\prime \prime}, \gamma^{\prime \prime}\right)$ are in general non-linear functions of the angles $(\alpha, \beta, \gamma)$ and $\left(\alpha^{\prime}, \beta^{\prime}, \gamma^{\prime}\right)$. The rotation matrices $D_{m m^{\prime}}^{(l)}(\alpha, \beta, \gamma)$ provide a $(2 l+1) \times(2 l+1)$-dimensional matrix representation of the rotation group.

Multiplying the defining expansion (5.150) by the inverse rotation matrix and using the fact that these are unitary, we obtain the corresponding expression

$$
\begin{equation*}
Y_{l m}(\theta, \phi)=\sum_{m^{\prime}=-l}^{l} D_{m m^{\prime}}^{(l) *}(\alpha, \beta, \gamma) Y_{l m^{\prime}}\left(\theta^{\prime}, \phi^{\prime}\right) \tag{5.156}
\end{equation*}
$$

Now we consider the special rotation where the new angles $\theta^{\prime}=\phi^{\prime}=0$. This is obtained when the axis system is rotated by $\alpha=\phi$ and $\beta=\theta$ as is seen from Fig.5.4 showing the situation the $x y$-plane. In this special case (5.156) gives

$$
Y_{l m}(\beta, \alpha)=\sum_{m^{\prime}=-l}^{l} D_{m m^{\prime}}^{(l) *}(\alpha, \beta, \gamma) Y_{l m^{\prime}}(0,0)
$$



Figure 5.7: Passive rotation by $\alpha$ in the $x y$-plane.

But from the previous section we know that the spherical harmonics $Y_{l m}$ vanish in the forward direction $\theta=0$ except for $m=0$. More accurately we have

$$
Y_{l m^{\prime}}(0,0)=\sqrt{\frac{2 l+1}{4 \pi}} \delta_{m^{\prime} 0}
$$

from (5.105). Inserted in the above sum, we then have the useful result

$$
\begin{equation*}
Y_{l m}(\beta, \alpha)=\sqrt{\frac{2 l+1}{4 \pi}} D_{m 0}^{(l) *}(\alpha, \beta, \gamma) \tag{5.157}
\end{equation*}
$$

This is seen when we now take the $m=0$ component of the expansion (5.150),

$$
Y_{l 0}\left(\theta^{\prime}, \phi^{\prime}\right)=\sum_{m^{\prime}=-l}^{l} Y_{l m^{\prime}}(\theta, \phi) D_{m^{\prime} 0}^{(l)}(\alpha, \beta, \gamma)
$$

Now using our result for $D_{m 0}^{(l)}$ and expressing the left-hand side in terms of a Legendre polynomial by (5.105), we have

$$
\begin{equation*}
P_{l}\left(\cos \theta^{\prime}\right)=\frac{4 \pi}{2 l+1} \sum_{m=-l}^{l} Y_{l m}(\theta, \phi) Y_{l m}^{*}(\beta, \alpha) \tag{5.158}
\end{equation*}
$$

This special relation is called the addition theorem for spherical harmonics. An equivalent form is obtained by introducing the two unit vectors $\mathbf{n}$ and $\mathbf{k}$ in the directions $(\theta, \phi)$ and $(\beta, \alpha)$ respectively. Then it can be written as

$$
\begin{equation*}
P_{l}(\mathbf{n} \cdot \mathbf{k})=\frac{4 \pi}{2 l+1} \sum_{m=-l}^{l} Y_{l m}(\mathbf{n}) Y_{l m}^{*}(\mathbf{k}) \tag{5.159}
\end{equation*}
$$

where $\cos \theta^{\prime}=\mathbf{n} \cdot \mathbf{k}=\cos \theta \cos \beta+\sin \theta \sin \beta \cos (\phi-\alpha)$. For $\theta^{\prime}=0$ is reduces to a completeness sum for spherical harmonics.

With the Euler parametrization the general rotation matrix element (5.151) becomes

$$
\begin{align*}
D_{m m^{\prime}}^{(l)}(\alpha, \beta, \gamma) & =\langle l, m| e^{-i \alpha \widehat{L}_{z} / \hbar} e^{-i \beta \widehat{L}_{y} / \hbar} e^{-i \gamma \widehat{L}_{z} / \hbar}\left|l, m^{\prime}\right\rangle  \tag{5.160}\\
& =e^{-i\left(m \alpha+m^{\prime} \gamma\right)} d_{m m^{\prime}}^{(l)}(\beta) \tag{5.161}
\end{align*}
$$

where

$$
\begin{equation*}
d_{m m^{\prime}}^{(l)}(\beta)=\langle l, m| e^{-i \beta \widehat{L}_{y} / \hbar}\left|l, m^{\prime}\right\rangle \tag{5.162}
\end{equation*}
$$

is the reduced rotation matrix. For the lowest spins they are easily found. The simplest case is $\operatorname{spin} S=1 / 2$. From (5.88) we then have

$$
e^{-i \beta \widehat{S}_{y} / \hbar}=e^{-i \sigma_{y} \beta / 2}=\cos \frac{\beta}{2}-i \sigma_{y} \sin \frac{\beta}{2}
$$

so that

$$
d^{(1 / 2)}(\beta)=\left(\begin{array}{rr}
\cos \beta / 2 & -\sin \beta / 2  \tag{5.163}\\
\sin \beta / 2 & \cos \beta / 2
\end{array}\right)
$$

With $\beta=\pi / 2$ we recover the special result (5.89).
For spin $S=1$ we take the generator $L_{y}$ from (5.90). As noted there, it has the property that $\left(L_{y} / \hbar\right)^{3}=L_{y} / \hbar$. Expanding and combining even and odd terms, one can then write

$$
\begin{equation*}
e^{-i \beta \widehat{L}_{y} / \hbar}=1-i \sin \beta\left(L_{y} / \hbar\right)-(1-\cos \beta)\left(L_{y} / \hbar\right)^{2} \tag{5.164}
\end{equation*}
$$

Now we have

$$
\left(\frac{L_{y}}{\hbar}\right)^{2}=\frac{1}{2}\left(\begin{array}{rrr}
1 & 0 & -1 \\
0 & 2 & 0 \\
-1 & 0 & 1
\end{array}\right)
$$

and thus one obtains the rotation matrix

$$
d^{(1)}(\beta)=\left(\begin{array}{ccc}
(1+\cos \beta) / 2 & -\sin \beta / \sqrt{2} & (1-\cos \beta) / 2  \tag{5.165}\\
\sin \beta / \sqrt{2} & \cos \beta & -\sin \beta / \sqrt{2} \\
(1-\cos \beta) / 2 & \sin \beta / \sqrt{2} & (1+\cos \beta) / 2
\end{array}\right)
$$

Wigner has derived a compact formula which allows the calculation of all higher spin rotation matrices.

In the same way as the spherical harmonics are eigenfunctions for the rigid rotator, these new rotation functions give the eigenfunctions for a general rigid body in rotation. That is the only motion it can have in its own center-of-mass system. To see this, we attach the rotated axis system $\left(x^{\prime}, y^{\prime}, z^{\prime}\right)$ to the body. The orientation of the body will then at any instant be given by the three Euler angles.

From classical mechanics we know that the Hamiltonian of such a rotating body is

$$
\begin{equation*}
H=\frac{L_{1}^{2}}{2 I_{1}}+\frac{L_{2}^{2}}{2 I_{2}}+\frac{L_{3}^{2}}{2 I_{3}} \tag{5.166}
\end{equation*}
$$

where $I_{1}, I_{2}$ and $I_{3}$ are the three principal moments of inertia. Similarly, $L_{1}, L_{2}$ and $L_{3}$ are the components of the conserved angular momentum vector $\mathbf{L}$ along the same axes, i.e. along the unit basis vectors $\left(\mathbf{e}_{1}, \mathbf{e}_{2}, \mathbf{e}_{3}\right)$ of the body-fixed coordinate system $\left(x^{\prime}, y^{\prime}, z^{\prime}\right)$.

In a quantum description of this motion described by the Euler angles, the components of the angular momentum become differential operators in these variables. We can find them from the different rotation vectors involved in each seprate rotation. The first rotation by $\alpha$ was around the laboratory-fixed $z$ axis so we have $\mathbf{n}_{\alpha}=\mathbf{e}_{z}$. From the defining expression (5.49) it then follows

$$
\begin{equation*}
L_{z}=-i \hbar \frac{\partial}{\partial \alpha} \tag{5.167}
\end{equation*}
$$

The second rotation by $\beta$ was around the $y^{\prime \prime}$-axis which is in the $x y$-plane an angle $\alpha$ away from the original $y$-axis. Thus $\mathbf{n}_{\beta}=-\sin \alpha \mathbf{e}_{x}+\cos \alpha \mathbf{e}_{y}$ giving

$$
-i \hbar \frac{\partial}{\partial \beta}=-\sin \alpha L_{x}+\cos \alpha L_{y}
$$

After a little more geometry we find the third rotation of $\gamma$ around the $z^{\prime}$ axis to be in the direction $\mathbf{n}_{\gamma}=\sin \beta \cos \alpha \mathbf{e}_{x}+\sin \beta \sin \alpha \mathbf{e}_{y}+\cos \beta \mathbf{e}_{z}$. We therefore have, again from (5.49),

$$
-i \hbar \frac{\partial}{\partial \gamma}=\sin \beta \cos \alpha L_{x}+\sin \beta \sin \alpha L_{y}+\cos \beta L_{z}
$$

Together with (5.167) we can now solve for $L_{x}$ and $L_{y}$ with the results

$$
\begin{align*}
L_{x} & =i \hbar \cos \alpha\left(\cot \beta \frac{\partial}{\partial \alpha}-\frac{1}{\sin \beta} \frac{\partial}{\partial \gamma}\right)+i \hbar \sin \alpha \frac{\partial}{\partial \beta}  \tag{5.168}\\
L_{y} & =i \hbar \sin \alpha\left(\cot \beta \frac{\partial}{\partial \alpha}-\frac{1}{\sin \beta} \frac{\partial}{\partial \gamma}\right)-i \hbar \cos \alpha \frac{\partial}{\partial \beta} \tag{5.169}
\end{align*}
$$

Some more work then verifis that $\left[L_{x}, L_{y}\right]=i \hbar L_{z}$ holds in this parametrization. Since the energy of the body is independent of its orientation with respect to the laboratoryfixed coordinate system $(x, y, z)$, both the eigenvalues of

$$
\begin{equation*}
\mathbf{L}^{2}=-\hbar^{2}\left[\frac{\partial^{2}}{\partial \beta^{2}}+\cot \beta \frac{\partial}{\partial \beta}+\frac{1}{\sin ^{2} \beta}\left(\frac{\partial^{2}}{\partial \alpha^{2}}+\frac{\partial^{2}}{\partial \gamma^{2}}-2 \cos \beta \frac{\partial^{2}}{\partial \alpha \partial \gamma}\right)\right] \tag{5.170}
\end{equation*}
$$

and $L_{z}$ in (5.167) are conserved quantum numbers, corresponding to the magnitude and direction of $\mathbf{L}$ as seen in the laboratory system.
But what we need in order to find the quantized energy from the Hamiltonian (5.166) are the components of the angular momentum in the body-fixed system of axes, i.e. $L_{i}=\mathbf{e}_{i} \cdot \mathbf{L}$ in the three directions $i=(1,2,3)$. Similar geometric considerations as above then give $\mathbf{n}_{\gamma}=\mathbf{e}_{3}$ and $\mathbf{n}_{\beta}=\sin \gamma \mathbf{e}_{1}+\cos \gamma \mathbf{e}_{2}$. Combined with the more cumbersome $\mathbf{n}_{\alpha}=-\sin \beta \cos \gamma \mathbf{e}_{1}+\sin \beta \sin \gamma \mathbf{e}_{2}+\cos \beta \mathbf{e}_{3}$ we then have the equivalent, linear relations
between the generators

$$
\begin{aligned}
-i \hbar \frac{\partial}{\partial \gamma} & =L_{3} \\
-i \hbar \frac{\partial}{\partial \beta} & =\sin \gamma L_{1}+\cos \gamma L_{2} \\
-i \hbar \frac{\partial}{\partial \alpha} & =-\sin \beta \cos \gamma L_{1}+\sin \beta \sin \gamma L_{2}+\cos \beta L_{3}
\end{aligned}
$$

Solving for $L_{1}$ and $L_{2}$ now gives

$$
\begin{align*}
L_{1} & =-i \hbar \cos \gamma\left(\cot \beta \frac{\partial}{\partial \gamma}-\frac{1}{\sin \beta} \frac{\partial}{\partial \alpha}\right)-i \hbar \sin \gamma \frac{\partial}{\partial \beta}  \tag{5.171}\\
L_{2} & =+i \hbar \sin \gamma\left(\cot \beta \frac{\partial}{\partial \gamma}-\frac{1}{\sin \beta} \frac{\partial}{\partial \alpha}\right)-i \hbar \cos \gamma \frac{\partial}{\partial \beta} \tag{5.172}
\end{align*}
$$

The total angular momentum operator $\mathbf{L}^{2}=L_{1}^{2}+L_{2}^{2}+L_{3}^{2}$ is as expected the same as in (5.170). But when we calculate the commutators between these last components, we find the same result but now with opposite sign,

$$
\begin{equation*}
\left[L_{1}, L_{2}\right]=\hbar^{2}\left(\cot ^{2} \beta+\frac{\partial}{\partial \beta} \cot \beta\right) \frac{\partial}{\partial \gamma}=-\hbar^{2} \frac{\partial}{\partial \gamma}=-i \hbar L_{3} \tag{5.173}
\end{equation*}
$$

It is a bit surprising, but is basically due to the fact that rotations from the body-fixed frame seem to take place in the opposite direction compared with observations in the laboratory-fixed frame. Also it should be noted that all these body-fixed components of the angular momentum commute with the components ( $L_{x}, L_{y}, L_{z}$ ) in the laboratory system.

We can now write out the Hamiltonian (5.166) in terms of these operators. But instead of solving the resulting differential eigenvalue equation, which is very difficult, we can make use of the rotation functions we have found. If the body has an orientation given by three Euler angles $(\phi, \theta, \chi)$, the wavefunction will be of the form $\Psi_{l m}(\phi, \theta, \chi)$. Here the subscripts indicate the two conserved quantum numbers corresponding to the total angular momentum and its component along the $z$-axis.

If we now perform a passive rotation $(\alpha, \beta, \gamma)$ of the coordinate system, the wavefunction will change into

$$
\begin{equation*}
\Psi_{l m}\left(\phi^{\prime}, \theta^{\prime}, \chi^{\prime}\right)=\sum_{k=-l}^{l} \Psi_{l k}(\phi, \theta, \chi) D_{k m}^{(l)}(\alpha, \beta, \gamma) \tag{5.174}
\end{equation*}
$$

in complete analogy with (5.150). Following the same line of reasoning, we again consider the special rotation $\alpha=\phi, \beta=\theta$ and $\chi=\gamma$ resulting in $\phi^{\prime}=\theta^{\prime}=\chi^{\prime}=0$. Multiplying the above equation with the rotation matrix for the inverse rotation and making use of its unitary property, we then have

$$
\begin{equation*}
\Psi_{l m}(\alpha, \beta, \gamma)=\sum_{k=-l}^{l} C_{l k} D_{m k}^{(l) *}(\alpha, \beta, \gamma) \tag{5.175}
\end{equation*}
$$

where the constants $C_{l k}=\Psi_{l k}(0,0,0)$. This result corresponds to (5.157) for the spherical harmonics. Thus we see that the rotation functions provide a basis for the eigenfunctions of a rotating, rigid body.

A symmetric top corresponds to such a body with a symmetry axis which we take to be the $z^{\prime}$ or 3 -axis. In the Hamiltonian (5.166) the two moments of inertia $I_{1}=I_{2}$ and it thus takes the form

$$
\begin{equation*}
H=\frac{1}{2 I_{1}} \mathbf{L}^{2}+\left(\frac{1}{2 I_{3}}-\frac{1}{2 I_{1}}\right) L_{3}^{2} \tag{5.176}
\end{equation*}
$$

Since an eigenfunction can at most pick up a phase factor under rotations around this axis, it must also be eigenstates of $L_{3}$ and therefore correspond to just one term in the sum (5.176). We then have the nice result that the eigenfunctions for the symmetric top are simply the complex rotation matrices,

$$
\begin{equation*}
\mathcal{Y}_{m k}^{l}(\alpha, \beta, \gamma)=N_{l} D_{m k}^{(l) *}(\alpha, \beta, \gamma) \tag{5.177}
\end{equation*}
$$

The normalization constant in front turns out to be $N_{l}=\sqrt{l+1 / 2} / 2 \pi$. These functions are obviously also eigenfunctions of $L_{z}$ and $L_{3}$

$$
L_{z} \mathcal{Y}_{m k}^{l}=m \hbar \mathcal{Y}_{m k}^{l}, \quad L_{3} \mathcal{Y}_{m k}^{l}=k \hbar \mathcal{Y}_{m k}^{l}
$$

as follows from the expression (5.167) for $L_{z}$ and the corresponding expression for $L_{3}$. The ladder operators $L_{ \pm}=L_{x} \pm i L_{y}$ will raise or lower the first index as before. Also $L_{ \pm}^{\prime}=L_{1} \pm i L_{2}$ act as raising and lowering operators, but on the last index of the wavefunction. In fact, due to the minus sign in (5.173), $L_{-}^{\prime}$ will act as a raising operator and $L_{+}^{\prime}$ as a lowering operator.

For the corresponding eigenvalues of the Hamiltonian (5.176) we then have

$$
\begin{equation*}
E_{l k}=\frac{\hbar^{2}}{2 I_{1}} l(l+1)+\left(\frac{1}{2 I_{3}}-\frac{1}{2 I_{1}}\right) \hbar^{2} k^{2} \tag{5.178}
\end{equation*}
$$

Each energy level specified by the quantum numbers $(l, k)$ has a degeneracy of $2 l+1$ corresponding to the different quantized directions of the angular momentum in the laboratory system. An example of such a system would be a rotating ammonium molecule $\mathrm{NH}_{3}$ where the 3 -axis goes through the nitrogen atom and normal to the plane with hydrogen atoms.

In the special case of a spherical top when the three principal moments are the same, $I_{1}=I_{2}=I_{3}$, the energies are independent of the quantum number $k$. As a consequence, the degeneracy of each level is increased by a factor of $2 l+1$ corresponding now to the different quantized directions of the angular momentum in the body-fixed refererence system. The total degeneracy is therefore $(2 l+1)^{2}$. A rotating methane $\mathrm{CH}_{4}$ molecule, where the carbon atom is symmetrically surounded by four hydrogen atoms, exhibits such a rotational spectrum.

## Chapter 6

## Approximation schemes

The harmonic oscillator is one of the few systems in quantum mechanics which allows us to find exact solutions both for the eigenvalues and eigenstates. For essentially all other systems it is necessary to make use of approximation methods in order to calculate these quantities from a given Hamiltonian. There are basically two different classes of such methods. The first is based on the variational principle and is especially suited for investigations of the ground state and the lowest excited states. In the other scheme one assumes that the real system is is approximately equal to another system which allows exact solutions. Then using perturbation theory one can systematically calculate corrections to this lowest and often crude approximation. This method can be applied to any state and makes it possible to calculate any quantity as accurately as is needed, at least in principle.

### 6.1 The variational method

Let the system under consideration be described by a Hamiltonian operator $\widehat{H}$. It has eigenvalues $E_{n}$ and corresponding eigenstates $\left|E_{N}\right\rangle$ satisfying

$$
\begin{equation*}
\widehat{H}\left|E_{n}\right\rangle=E_{n}\left|E_{n}\right\rangle \tag{6.1}
\end{equation*}
$$

The problem is that in most cases this equation cannot be solved exactly, we only know that in principle there are solutions to it.
Consider now some general state $|\psi\rangle$ which always can be expanded in the basis provided by the unknown eigenstates $\left|E_{n}\right\rangle$ as

$$
\begin{equation*}
|\psi\rangle=\sum_{n=0}^{\infty} C_{n}\left|E_{n}\right\rangle \tag{6.2}
\end{equation*}
$$

with components $C_{n}$. We assume that the states are labelled so that the ground state corresponds to $n=0$, the first excited has $n=1$ and so on. If the state is normalized to one, the components will satisfy

$$
\begin{equation*}
\sum_{n=0}^{\infty}\left|C_{n}\right|^{2}=1 \tag{6.3}
\end{equation*}
$$

since the eigenstates $\left|E_{n}\right\rangle$ are orthonormalized. If we now the calculate the matrix element $E_{\psi}=\langle\psi| \widehat{H}|\psi\rangle$, we obtain

$$
E_{\psi}=\sum_{n=0}^{\infty} C_{n}\langle\psi| \widehat{H}\left|E_{n}\right\rangle=\sum_{n=0}^{\infty}\left|C_{n}\right|^{2} E_{n}
$$

On the right-hand side we can now replace all the energies $E_{n}$ with the lowest energy $E_{0}$. Since $E_{n} \geq E_{0}$ for all $n>0$, we then find the fundamental inequality

$$
E_{\psi} \geq E_{0} \sum_{n=0}^{\infty}\left|C_{n}\right|^{2}=E_{0}
$$

using the normalization (6.3). So for any choice of the state $|\psi\rangle$ we will always find a result which is above the true ground state energy $E_{0}$. The exact value is obtained only by constructing this trial state vector to be the correct ground state, something which is generally impossible.

Any choice for the state $|\psi\rangle$ will involve one or more unknown parameters $a_{i}$. These can now be determined by requiring that the resulting energy $E_{\psi}$ is as small as possible, i.e. we have

$$
\begin{equation*}
\frac{\partial E_{\psi}}{\partial a_{i}}=0 \tag{6.4}
\end{equation*}
$$

With this choice of parameters we then have the best upper bound for the true ground state energy $E_{0}$.
In order to illustrate the use of this variational method, consider a particle with mass $m$ moving in one dimension described by the Hamiltonian operator

$$
\widehat{H}=\frac{1}{2 m} \widehat{p}^{2}+\widehat{V}
$$

where the potential energy is

$$
V(x)= \begin{cases}\infty, & x<0  \tag{6.5}\\ k x, & x \geq 0\end{cases}
$$

in the coordinate representation. The ground state wave function must then be zero for $x=0$ since the probability to find the particle for $x \leq 0$ must be zero. Likewise it must go to zero when $x \rightarrow \infty$. We therefore choose the trial wavefunction $\psi(x)=\langle x \mid \psi\rangle$ to be

$$
\psi(x)=2 a^{3 / 2} x e^{-a x}
$$

where the prefactor is determined by the normalization

$$
\int_{0}^{\infty} d x|\psi(x)|^{2}=1
$$

The unknown quantity $a$ can then be taken as the variational parameter. Here and in similar integrals we make use of Euler's gamma-function defined by

$$
\begin{equation*}
\int_{0}^{\infty} d x x^{s} e^{-x}=\Gamma(s+1)=s \Gamma(s) \tag{6.6}
\end{equation*}
$$

When $s=n$ is a postive integer, then $\Gamma(n+1)=n!$ with $0!=1$. In addition it is useful to remember that $\Gamma(1 / 2)=\sqrt{\pi}$ as follows from the Gaussian integral

$$
\begin{equation*}
\int_{0}^{\infty} d x e^{-x^{2}}=\frac{1}{2} \sqrt{\pi} \tag{6.7}
\end{equation*}
$$

With the above choice for the trial wave function, we then find the matrix kinetic element

$$
\begin{aligned}
\langle\psi| \frac{\hat{p}^{2}}{2 m}|\psi\rangle & =4 a^{3} \int_{0}^{\infty} d x x e^{-a x}\left(-\frac{\hbar^{2}}{2 m} \frac{d^{2}}{d x^{2}}\right) x e^{-a x} \\
& =-4 a^{3} \frac{\hbar^{2}}{2 m} \int_{0}^{\infty} d x\left(-2 a x+a^{2} x^{2}\right) e^{-2 a x}=\frac{\hbar^{2}}{2 m} a^{2}
\end{aligned}
$$

and similarly for the potential energy

$$
\langle\psi| \widehat{V}|\psi\rangle=4 a^{3} k \int_{0}^{\infty} d x x^{3} e^{-2 a x}=\frac{3 k}{2 a}
$$

We thus have

$$
E_{\psi}=\frac{\hbar^{2}}{2 m} a^{2}+\frac{3 k}{2 a}
$$

which has a minimum for $a=\left(3 k m / 2 \hbar^{2}\right)^{1 / 3}$. Calculating the corresponding value of $E_{\psi}$, we find that the ground state energy will therefore satisfy the bound

$$
\begin{equation*}
E_{0} \leq \frac{3}{4} 6^{2 / 3}\left(\frac{\hbar^{2} k^{2}}{2 m}\right)^{1 / 3} \tag{6.8}
\end{equation*}
$$

The numerical prefactor $3 \cdot 6^{2 / 3} / 4=2.48$ while an exact solution using Airy functions gives 2.34. Our variational result is therefore within $6 \%$ of the correct value. We can get much closer to this value by choosing a more complex trial function with more parameters but with the same general properties as we imposed on our simple choice above.

It is possible to use the variational method to estimate also the energy of the first excited state with a similar accuracy. We must then use a trial wave function which is orthogonal to the true ground state. This is achieved by taking

$$
\begin{equation*}
|\psi\rangle=\sum_{n=1}^{\infty} C_{n}\left|E_{n}\right\rangle \tag{6.9}
\end{equation*}
$$

instead of (6.2). Since now we have the normalization $\sum_{n=1}^{\infty}\left|C_{n}\right|^{2}=1$, we get

$$
\langle\psi| \widehat{H}|\psi\rangle=\sum_{n=1}^{\infty}\left|C_{n}\right|^{2} E_{n} \geq E_{1}
$$

which can then be minimized by the variational method. In practice the general form of trial wave function for the first excited state is often determined by having a symmetry like parity different from what the ground state has.

### 6.2 Rayleigh-Ritz scheme

The above variational method is useful for accurate calculations of the lowest two states of a quantum system. An alternative scheme is due to Rayleigh and Ritz and was developed before quantum mechanics. It can in many cases be used to determine also the higher excited states in a systematic way and especially with the use of modern computers.

It is based on the assumption that one can find a similar system with the same variables but simpler dynamics so that it has an exact and known solution. This model system and the real system under consideration will then have the same Hilbert spaces. In the example with the wedge potential (6.5) in the previous section, the simpler system could then have a harmonic potential in the region $x>0$ instead of the linear potential.
Denoting the Hamiltonian of the simpler or ideal system by $\widehat{H}_{0}$, we therefore assume that the eigenvalue problem $\widehat{H}_{0}|n\rangle=E_{n}^{0}|n\rangle$ is solved. The orthonormalized eigenstates $|n\rangle$ thus form a complete set. A general state $|\psi\rangle$ of the real system with Hamiltonian $\widehat{H}$ can therefore be expanded as

$$
\begin{equation*}
|\psi\rangle=\sum_{n=0}^{\infty} a_{n}|n\rangle \tag{6.10}
\end{equation*}
$$

where the components $a_{n}=\langle n \mid \psi\rangle$ are unknown and will be considered as variational parameters in the following. Since this vector is normalized to one, they satisfy the constraint

$$
\begin{equation*}
\langle\psi \mid \psi\rangle=\sum_{m=0}^{\infty} a_{m}^{*} a_{m}=1 \tag{6.11}
\end{equation*}
$$

Calculating now the expectation value $E_{\psi}=\langle\psi| \widehat{H}|\psi\rangle$, one obtains

$$
E_{\psi}=\sum_{m, n=0}^{\infty} a_{m}^{*} a_{n} H_{m n}
$$

where the matrix element $H_{m n}=\langle m| \widehat{H}|n\rangle$. These can now all be calculated since the states $|n\rangle$ are known. We can now minimize $E_{\psi}$ under variations of the parameters $a_{n}$. Since these are not all free, but are restricted by (6.11), we introduce a Lagrange multiplier $E$ and minimize instead

$$
\begin{align*}
F_{\psi} & =E_{\psi}-E\langle\psi \mid \psi\rangle \\
& =\sum_{m, n=0}^{\infty} a_{m}^{*}\left(H_{m n}-E \delta_{m n}\right) a_{n} \tag{6.12}
\end{align*}
$$

Since the parameters $a_{n}$ in general are complex, we should now minimize this with respect to both the real an imaginary parts. But this is equivalent to minimizing with respect to $a_{n}$ and $a_{n}^{*}$. Thus we must have

$$
\frac{\partial F_{\psi}}{\partial a_{m}^{*}}=\sum_{n=0}^{\infty}\left(H_{m n}-E \delta_{m n}\right) a_{n}=0
$$



Figure 6.1: Energy-levels obtained in the Rayleigh-Ritz scheme are upper bounds for the true eigenvalues obtained for $N=\infty$.
and the corresponding compex conjugated equation which doesn't give anything new. Writing out this matrix equation, we find an infinite set of linear and homogenous equations for the unknown parameters $a_{n}$. Non-trivial solutions are then found only when

$$
\begin{equation*}
\operatorname{det}\left(H_{m n}-E \delta_{m n}\right)=0 \tag{6.13}
\end{equation*}
$$

The zeros of this $\infty \times \infty$ determinant are then the energy eigenvalues $E_{n}$ we are seeking. And if we could actually perform the calculation, these would also be the exact eigenvalues.
In practice we can only include a finite number $N$ of the lowest eigenstates $|n\rangle$ so that the determinant (6.13) has dimension $N \times N$. Writing it out, the characteristic polynomial in $E$ will have $N$ zeros $E_{0}^{(N)}, E_{1}^{(N)}, E_{2}^{(N)}, \ldots, E_{N}^{(N)}$ which will be upper limits to the corresponding true eigenvalues. In fact, the Norwegian physicists Hylleraas and Undheim showed in th 30 's that these approximate values will all satisfy the bounds $E_{k}^{(n)} \geq E_{k}^{(n+1)}$ as indicated in the figure. The larger the number $N$ of states included in the calculation, the closer the approximation will be.

### 6.3 Static perturbations

It should be pretty obvious that the Rayleigh-Ritz method will converge faster towards the correct eigenvalues the closer the reference system described by $\widehat{H}_{0}$ is to the real
system described by the Hamiltonian $\widehat{H}$. We can then write

$$
\begin{equation*}
\widehat{H}=\widehat{H}_{0}+\widehat{V} \tag{6.14}
\end{equation*}
$$

where the effect of the operator $\widehat{V}$ is in some sense small. It is said to perturb the dynamics described by $\widehat{H}_{0}$ which we again will assume can be calculated exactly. Introducing here a slightly different notation, we denote these eigenstates by $|n\rangle$ so that $\widehat{H}_{0}|n\rangle=E_{n}|n\rangle$. The eigenstates of the full Hamiltonian $\widehat{H}$ we denote by $\left|n^{\prime}\right\rangle$ with the corresponding eigenvalues $E_{n}^{\prime}=E_{n^{\prime}}$ satisfying

$$
\begin{equation*}
\widehat{H}\left|n^{\prime}\right\rangle=E_{n}^{\prime}\left|n^{\prime}\right\rangle \tag{6.15}
\end{equation*}
$$

Since the effect of the perturbation $\widehat{V}$ is assumed to be small, the differences $\Delta E_{n}=$ $E_{n}^{\prime}-E_{n}$ betweeen corresponding eigenvalues will ordinarily also be small. This is schematically illustrated in Fig.6.2. Assuming that the perturbation $\widehat{V}$ is independent


Figure 6.2: Unperturbed and perturbed energy levels.
of time, we will now calculate the first corrections of these energy shifts in a systematic way appropriately called static perturbation theory.
Since the eigenstates of the 'free' system described by $\widehat{H}_{0}$ form a complete set, we can expand an 'interacting' state $\left|n^{\prime}\right\rangle$ as

$$
\begin{equation*}
\left|n^{\prime}\right\rangle=\sum_{m}|m\rangle\left\langle m \mid n^{\prime}\right\rangle=\sum_{m} a_{m n^{\prime}}|m\rangle \tag{6.16}
\end{equation*}
$$

where the coefficient $a_{m n^{\prime}}=\left\langle m \mid n^{\prime}\right\rangle$. We expect that $a_{n n^{\prime}} \rightarrow 1$ and otherwise when $m \neq n$ that $a_{m n^{\prime}} \rightarrow 0$ as $V \rightarrow 0$.

Letting the Hamiltonian (6.14) act on state $\left|n^{\prime}\right\rangle$ and projecting the resulting equation onto $\langle m|$, one has

$$
\langle m| \widehat{H}-\widehat{H}_{0}\left|n^{\prime}\right\rangle=\langle m| \widehat{V}\left|n^{\prime}\right\rangle
$$

Since these vectors are eigenstates of $\widehat{H}$ and $\widehat{H}_{0}$ respectively, we obtain

$$
\begin{equation*}
\langle m| E_{n^{\prime}}-E_{m}\left|n^{\prime}\right\rangle=\left(E_{n^{\prime}}-E_{m}\right)\left\langle m \mid n^{\prime}\right\rangle=\langle m| \widehat{V}\left|n^{\prime}\right\rangle \tag{6.17}
\end{equation*}
$$

When $m \neq n$ we therefore have for the expansion coefficients

$$
\begin{equation*}
a_{m n^{\prime}}=\frac{\langle m| \widehat{V}\left|n^{\prime}\right\rangle}{E_{n^{\prime}}-E_{m}} \tag{6.18}
\end{equation*}
$$

On the other hand, in the special case $m=n$ we get for the sought-for energy difference

$$
\begin{equation*}
E_{n}^{\prime}-E_{n}=\Delta E_{n}=\frac{\langle n| \hat{V}\left|n^{\prime}\right\rangle}{\left\langle n \mid n^{\prime}\right\rangle} \tag{6.19}
\end{equation*}
$$

It is now convenient to normalize the state $\left|n^{\prime}\right\rangle$ so that $a_{n n^{\prime}}=\left\langle n \mid n^{\prime}\right\rangle=1$. The perturbed energy is then given by the simple expression

$$
\begin{equation*}
E_{n}^{\prime}=E_{n}+\langle n| \widehat{V}\left|n^{\prime}\right\rangle \tag{6.20}
\end{equation*}
$$

which we can make good use of.
In the lowest approximation, we can write that the perturbed state $\left|n^{\prime}\right\rangle=|n\rangle+\ldots$ where the dots are higher order corrections. We then get the result

$$
\begin{equation*}
E_{n}^{\prime}=E_{n}+\langle n| \widehat{V}|n\rangle \tag{6.21}
\end{equation*}
$$

which is always simple to use and gives the dominant effect. Since in this lowest approximation the level shift is proportional to $V$, it is said to be correct to the first order in the perturbation.
Some times the first-order shift is zero or we want better accuracy. We must then calculate to next order in the perturbation. From the general formula (6.20) we see that this can be obtained by including the first-order correction in the perturbed state $\left|n^{\prime}\right\rangle$. Using that $a_{n n^{\prime}}=1$, we can write the expansion (6.16) as

$$
\left|n^{\prime}\right\rangle=|n\rangle+\sum_{m \neq n} a_{m n^{\prime}}|m\rangle
$$

Here we can use (6.18) for the coefficients $a_{m n^{\prime}}$ when $m \neq n$. Since we now want the level shift to second order in $V$, we only need $a_{m n^{\prime}}$ to first order in $V$. We can therefore replace $E_{n^{\prime}}$ with $E_{n}$ and $\left|n^{\prime}\right\rangle$ with $|n\rangle$ in (6.18) giving

$$
\begin{equation*}
\left|n^{\prime}\right\rangle=|n\rangle+\sum_{m \neq n} \frac{\langle m| \hat{V}|n\rangle}{E_{n}-E_{m}}|m\rangle \tag{6.22}
\end{equation*}
$$

and therefore for the second-order level shift

$$
\begin{equation*}
E_{n}^{\prime}=E_{n}+\langle n| \widehat{V}|n\rangle+\sum_{m \neq n} \frac{|\langle m| \widehat{V}| n\rangle\left.\right|^{2}}{E_{n}-E_{m}} \tag{6.23}
\end{equation*}
$$

from (6.20). Higher order corrections can be calculated from the same perturbative expansion.

As a very simple example consider a 2-dimensional Hilbert space with

$$
H_{0}=\left(\begin{array}{rr}
0 & -3 \\
-3 & 8
\end{array}\right)
$$

and the perturbation

$$
V=\left(\begin{array}{rr}
0 & 1 \\
1 & -5
\end{array}\right)
$$

We easily find the two unperturbed eigenvalues $E_{1}=9$ and $E_{2}=-1$ with the corresponding normalized eigenvectors

$$
\psi_{1}=\sqrt{\frac{1}{10}}\binom{1}{-3}, \quad \psi_{2}=\sqrt{\frac{1}{10}}\binom{3}{1}
$$

To first order we need the matrix elements $\langle 1| \widehat{V}|1\rangle=-51 / 10$ and $\langle 2| \widehat{V}|2\rangle=1 / 10$ giving the perturbed energies $E_{1}^{\prime}=9-5.1=3.9$ and $E_{2}^{\prime}=-1+0.1=-0.9$. To second order we also need the matrix element $\langle 1| \widehat{V}|2\rangle=7 / 10$ which gives the improved eigenvalues

$$
E_{1}^{\prime}=3.9+\frac{(0.7)^{2}}{9+1}=3.9+0.05=3.95
$$

and

$$
E_{2}^{\prime}=-0.9+\frac{(0.7)^{2}}{-1-9}=-0.9-0.05=-0.95
$$

The second order correction is thus already pretty small. From (6.22) we can now find the corresponding, approximate eigenstates. Here we can also easily calculate the exact energies which are the eigenvalues of the matrix

$$
H=H_{0}+V=\left(\begin{array}{rr}
0 & -2  \tag{6.24}\\
-2 & 3
\end{array}\right)
$$

i.e. $E_{1}^{\prime}=4$ and $E_{2}^{\prime}=-1$. Our perturbative results are thus accurate to around $5 \%$.

### 6.4 Degeneracy

In the above we have assumed that the energy level under consideration is not degenerate. If that is the case, we must be a bit more careful. For instance, the result (6.18) for the expansion coefficients becomes meaningless if the unperturbed levels $E_{m}$ and $E_{n}$ are the same. A way out of the problem can be seen from (6.17). Since the left-hand side then would go to zero with degeneracy, we will only have consistency if the right-hand side also goes to zero, i.e. the non-diagonal matrix element $\langle m| \widehat{V}|n\rangle$ must also approach zero.

To make this consideration a bit more systematic, one can isolate the $g$ states degenerate with the state $|n\rangle$ under consideration. For this purpose we write the expansion (6.16) as

$$
\left|n^{\prime}\right\rangle=\sum_{s=1}^{g} a_{s}\left|n_{s}\right\rangle+\sum_{m \neq n} a_{m n^{\prime}}|m\rangle \equiv \sum_{s=1}^{g} a_{s}\left|n_{s}\right\rangle+\left|n_{\perp}\right\rangle
$$

where the state $\left|n_{\perp}\right\rangle$ is orthogonal to all the states $\left|n_{s}\right\rangle$. This state is now an eigenstate of $\widehat{H}_{0}+\widehat{V}$ with eigenvalue $E_{n}+\Delta E_{n}$. Writing out this condition, we get to 0th order

$$
\widehat{H}_{0} \sum_{s=1}^{g} a_{s}\left|n_{s}\right\rangle=E_{n} \sum_{s=1}^{g} a_{s}\left|n_{s}\right\rangle
$$

which is trivially satisfied. However, to 1st order we get the more interesting equation

$$
\left(\widehat{H}_{0}-E_{n}\right)\left|n_{\perp}\right\rangle+\left(\widehat{V}-\Delta E_{n}\right) \sum_{s=1}^{g} a_{s}\left|n_{s}\right\rangle=0
$$

Multiplication from the left with one of the degenerate states $\left\langle n_{r}\right|$ results in

$$
\left\langle n_{r}\right| \widehat{H}_{0}-E_{n}\left|n_{\perp}\right\rangle+\left\langle n_{r}\right| \widehat{V}-\Delta E_{n} \sum_{s=1}^{g} a_{s}\left|n_{s}\right\rangle=0
$$

But the first term here is zero since $\left\langle n_{r}\right| \widehat{H}_{0}=\left\langle n_{r}\right| E_{n}$. We are thus left with the second term which can be written as

$$
\begin{equation*}
\sum_{s=1}^{g}\left(V_{r s}-\Delta E_{n} \delta_{r s}\right) a_{s}=0 \tag{6.25}
\end{equation*}
$$

with the matrix elements $V_{r s}=\left\langle n_{r}\right| \hat{V}\left|n_{s}\right\rangle$. This set of $g$ linear and homogeneous


Figure 6.3: Degenerate energy level is split up by perturbation.
equations for the unknown coefficients $a_{s}$ has a solution only when the determinant

$$
\begin{equation*}
\operatorname{det}\left(V_{r s}-\Delta E_{n} \delta_{r s}\right)=0 \tag{6.26}
\end{equation*}
$$

It has $g$ roots which are the sought-for perturbations $\Delta E_{n}$.
Not all of these level shifts are necessarily different, some degeneracy can still remain. We see that even with degeneracy we can use non-degenerate perturbation theory if we choose a basis among the degenerate states so that the perturbation is diagonal.

Let us consider an example. An unperturbed rotor is given by the standard Hamiltonian $\widehat{H}=A \widehat{\mathbf{L}}^{2}$ where $A$ is a constant. The corresponding energy levels are $E_{\ell}=A \hbar^{2} \ell(\ell+1)$ where each level consists of $2 \ell+1$ degenerate states. We now place the rotor in a crystal field where the rotational invariance is broken and we take the corresponding perturbation to be $\widehat{V}=a \widehat{L}_{x}^{2}+b \widehat{L}_{y}^{2}+a \widehat{L}_{z}^{2}$. With a remaining rotational symmetry around the $z$-axis the parameters $a=b$. Now we want to see how this effects the energy levels of the system. Let us consider the level with $\ell=1$. Since it is 3 -fold degenerate, we must now calculate the matrix representation of the perturbation in the unperturbed basis. For the last term involving $\widehat{L}_{z}^{2}$, this is trivial. In the two first terms we can make use of the ladder operators $\widehat{L}_{ \pm}$and write $\widehat{L}_{x}^{2}=\left(\widehat{L}_{+}^{2}+\widehat{L}_{-}^{2}+\widehat{L}_{+} \widehat{L}_{-}+\widehat{L}_{-} \widehat{L}_{+}\right) / 4$ and correspondingly for $\widehat{L}_{y}^{2}$. For the matrix elements we can then make use of $\hat{L}_{-}|1,+1\rangle=2 \hbar^{2}|1,-1\rangle$ and correspondingly $\widehat{L}_{+}|1,-1\rangle=2 \hbar^{2}|1,+1\rangle$. The operators $\widehat{L}_{-} \widehat{L}_{+}$and $\widehat{L}_{+} \widehat{L}_{-}$will not change the $z$-component of the spin, i.e. will not change the states they act upon. We find that they just give $2 \hbar^{2}$ times the same state in this case. In this way we obtain for the perturbing matrix

$$
V=\hbar^{2}\left(\begin{array}{ccc}
\alpha & 0 & \gamma \\
0 & \beta & 0 \\
\gamma & 0 & \alpha
\end{array}\right)
$$

where the entries are $\alpha=c+(a+b) / 2, \beta=a+b$ and $\gamma=(a-b) / 2$. Remarkably enough, the eigenvalues of this matrix come out simply as $a+b, a+c$ and $b+c$. These are the amounts the three degenerate levels are shifted by the perturbation. With rotational symmetry around the $z$-axis, we see that two of the levels remain degenerate. This is as expected. With full rotation symmetry, i.e. $a=b=c$ all three levels remain degenerate, but are shifted by the common amount $2 a \hbar^{2}$. The perturbation is then simply $\widehat{V}=a \widehat{\mathbf{L}}^{2}$ which just equals $2 a \hbar^{2}$ when $\ell=1$.

### 6.5 Time-dependent perturbations

So far we have considered static perturbations. However, in many important physical situations the perturbing interaction depends on time so that the full Hamiltonian

$$
\begin{equation*}
\widehat{H}=\widehat{H}_{0}+\widehat{V}(t) \tag{6.27}
\end{equation*}
$$

will also depend on time. The evolution of the state vector will still be governed by the Schrödinger equation

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t}|\Psi, t\rangle=\widehat{H}|\Psi, t\rangle \tag{6.28}
\end{equation*}
$$

but will no longer be given by the simple time-evolution operator (2.35) in Chapter 2. Instead, we proceed as in Chapter 3 and expand the state vector in a complete and orthonormal set so that

$$
\begin{equation*}
|\Psi, t\rangle=\sum_{n} c_{n}(t)|n\rangle \tag{6.29}
\end{equation*}
$$

where now the expansion coefficients $c_{n}(t)=\langle n \mid \Psi, t\rangle$ are the components of the vector along the basis vector $|n\rangle$. The single equation (6.28) is then transformed into the coupled, first-order differential equations

$$
\begin{equation*}
i \hbar \dot{c}_{m}=\sum_{n} H_{m n} c_{n} \tag{6.30}
\end{equation*}
$$

exactly as in (3.2) with $H_{m n}=\langle m| \widehat{H}|n\rangle$. These equations are valid also when the Hamiltonian depends on time. Describing the time evolution of the system is now equivalent to calculating the amplitudes $c_{m}(t)$. They give us the probability $\left|c_{m}(t)\right|^{2}$ that the system is in the state $|n\rangle$ at time $t$.
When we now have the perturbative problem decribed by (6.27), it is natural to use the eigenstates of $\widehat{H}_{0}$ as the basis in the above expansion of the state vector. Then $\widehat{H}_{0}|n\rangle=$ $E_{n}|n\rangle$ where $E_{n}$ is the unperturbed energy of the system. The Hamiltonian matrix element in (6.30) then becomes $H_{m n}=E_{m} \delta_{m n}+V_{m n}(t)$ where $V_{m n}(t)=\langle m| \widehat{V}(t)|n\rangle$. In the resulting Schrödinger matrix equations

$$
i \hbar \dot{c}_{m}=E_{m} c_{m}+\sum_{n} V_{m n}(t) c_{n}
$$

the first term will dominate on the right-hand side since magnitude of the perturbation is assumed to be small compared to the unperturbed energies. We can separate out this dominant term by writing

$$
\begin{equation*}
c_{m}(t)=a_{m}(t) e^{-i E_{m} t / \hbar} \tag{6.31}
\end{equation*}
$$

Inserting now the time derivative

$$
i \hbar \dot{c}_{m}=\left(i \hbar \dot{a}_{m}+E_{m} a_{m}\right) e^{-i E_{m} t / \hbar}
$$

into the above equations of motion for the amplitudes $c_{m}$, we see that the energy term cancels out and we are left the differential equations

$$
\begin{equation*}
i \hbar \dot{a}_{m}=\sum_{n} V_{m n}(t) e^{i\left(E_{m}-E_{n}\right) t / \hbar} a_{n} \tag{6.32}
\end{equation*}
$$

for these reduced amplitudes. They still give the probability for the system to be found in state $|m\rangle$ since $\left|c_{m}(t)\right|^{2}=\left|a_{m}(t)\right|^{2}$ from (6.31).
So far we have not made any approximations. The coupled equations (6.32) are still exactly equivalent to the original Schrödinger equation and just as difficult to solve. But they are now ideally suited to describe a typical physical situation where the system is in some stationary state $|i\rangle$ until some time $t_{i}$ when a weak perturbation starts to act and we want to find out what happens to the system later. At the initial moment we know that the amplitudes are given by $a_{n}\left(t_{i}\right)=\delta_{n i}$, i.e. zero probability to be found in any other state than $|i\rangle$. When the perturbation then starts to act, all these amplitudes will change, but not so much since the perturbation is assumed to be small. We can then solve the set of equations (6.32) in a systematic way by inserting the 0th order solution $a_{n}(t)=\delta_{n i}$ on the right-hand side and then integrate to find the 1st order amplitudes. The amplitude which gives the probability for a quantum transition to a state $|f\rangle \neq|i\rangle$ at a time $t>t_{i}$ is therefore

$$
\begin{equation*}
a_{f}(t)=-\frac{i}{\hbar} \int_{t_{i}}^{t} d t^{\prime} V_{f i}\left(t^{\prime}\right) e^{i \omega_{f i} t^{\prime}} \tag{6.33}
\end{equation*}
$$

where the frequency $\omega_{f i}=\left(E_{f}-E_{i}\right) / \hbar$. This is the basic equation in first-order, timedependent perturbation theory. It is valid as long as the transition amplitude $\left|a_{f}\right| \ll 1$. Higher order corrections can similarly be derived by using this result on the right-hand side of (6.32). This would have given us the second order transition amplitudes and so on. We will not pursue that here.
As a simple example, consider the one-dimensional harmonic oscillator with the usual Hamiltonian

$$
\begin{equation*}
\widehat{H}_{0}=\frac{1}{2 m} \widehat{p}^{2}+\frac{1}{2} m \omega^{2} \widehat{x}^{2} \tag{6.34}
\end{equation*}
$$

perturbed by a small potential

$$
\begin{equation*}
\widehat{V}(t)=k \widehat{x} e^{-t^{2} / \tau^{2}} \tag{6.35}
\end{equation*}
$$

It describes a force which starts initially out to be zero, grows to a maximum value $k$ at time $t=0$ and then goes to zero again. The parameter $\tau$ gives the duration of this impuls. As intial condition we assume that the oscillator is in its ground state $|0\rangle$. Let us now calculate the amplitude $a_{f}$ to an excited state $|n\rangle$. Expessing the position operator $\widehat{x}$ in terms of ladder operators as in (4.68), we find for the transition matrix element

$$
V_{f i}=k \sqrt{\frac{\hbar}{2 m \omega}}\langle n| \widehat{a}^{\dagger}+\widehat{a}|0\rangle e^{-t^{2} / \tau^{2}}
$$

But $\widehat{a}|0\rangle=0$ by definition and $\widehat{a}^{\dagger}|0\rangle=|1\rangle$ so that it will be non-zero only for $n=1$. We thus get only transitions to the first excited state of the oscillator with $\omega_{f i}=\omega$. The transition amplitude (6.33) becomes

$$
a_{f}(t)=-\frac{i}{\hbar} k \sqrt{\frac{\hbar}{2 m \omega}} \int_{-\infty}^{\infty} d t e^{-t^{2} / \tau^{2}} e^{i \omega t}
$$

We can do the integral using

$$
\begin{equation*}
\int_{-\infty}^{\infty} d x e^{-\frac{1}{2} a x^{2}+b x}=\sqrt{\frac{2 \pi}{a}} e^{b^{2} / 2 a} \tag{6.36}
\end{equation*}
$$

even when the parameter $b$ is complex. It then gives

$$
\begin{equation*}
P(0 \rightarrow 1)=\left|a_{f}(\infty)\right|^{2}=\frac{\pi k^{2} \tau^{2}}{2 m \hbar \omega} e^{-\omega^{2} \tau^{2} / 2} \tag{6.37}
\end{equation*}
$$

for the transition probability.
The change in excitation number $|\Delta n|=1$ we find here is seen to result for any initial state. It is an example of a selection rule which says which final states can be reached for a given interaction. Had the interaction instead varied like $\widehat{V} \propto \widehat{x}^{2}$, we would have found the selection rules $|\Delta n|=2$ or $\Delta n=0$. In the last case it is said that we have an elastic transition since the initial and final states have the same energy.

### 6.6 Harmonic perturbations

In many real situations the time-dependency of the perturbation will be harmonic of the form

$$
\begin{equation*}
\widehat{V}(t)=2 \widehat{V} \cos \omega t=\widehat{V}\left(e^{-i \omega t}+e^{i \omega t}\right) \tag{6.38}
\end{equation*}
$$

It differs from the situations considered in the previous section in that it doesn't go away at very early or late times $t \rightarrow \pm \infty$. And we want to have unperturbed intial and final states in order to have a well-defined problem. This we can attain by assuming that the potential is slowly turned on at very early times and then slowly turned off again at very late times after the interaction is over. Mathematically, these adiabatic changes are achieved by modifying the perturbating potential to

$$
\begin{equation*}
\widehat{V}(t) \rightarrow \widehat{V^{\prime}}(t)=e^{-\epsilon|t|} \widehat{V}(t) \tag{6.39}
\end{equation*}
$$

where $\epsilon$ is a very small frequency. For times $|t| \gg 1 / \epsilon$ the perturbation is then zero as we want.

Both exponential terms will contribute to the transition amplitude (6.33). Let us consider the first term alone. If it is turned on at an early time $-T \ll-1 / \epsilon$ and turned off at a much later time $T \gg 1 / \epsilon$, we have from (6.33) the transition amplitude

$$
a_{f i}=-\frac{i}{\hbar} \int_{-T}^{T} d t V_{f i}(t) e^{i \omega_{f i} t-\epsilon|t|}=-\frac{i}{\hbar} V_{f i} \int_{-T}^{T} d t e^{i\left(\omega_{f i}-\omega\right) t-\epsilon|t|}
$$

From the last integral we have pulled out the time-independent matrix element $V_{f i}$. It is now well-defined and finite. Splitting it up,

$$
a_{f i} a_{f i}=-\frac{i}{\hbar} V_{f i}\left[\int_{-T}^{0} d t e^{i\left(\omega_{f i}-\omega-i \epsilon\right) t}+\int_{0}^{T} d t e^{i\left(\omega_{f i}-\omega+i \epsilon\right) t}\right]
$$

the two terms are easy to integrate,

$$
a_{f i}=-\frac{i}{\hbar} V_{f i}\left[\frac{1-e^{-i\left(\omega_{f i}-\omega-i \epsilon\right) T}}{i\left(\omega_{f i}-\omega-i \epsilon\right)}+\frac{e^{i\left(\omega_{f i}-\omega+i \epsilon\right) T}-1}{i\left(\omega_{f i}-\omega+i \epsilon\right)}\right]
$$

The exponentials in the numerators now disappear in the limit $T \rightarrow \infty$ and we are left with

$$
\begin{aligned}
a_{f i} & =-\frac{1}{\hbar} V_{f i}\left(\frac{1}{\omega_{f i}-\omega-i \epsilon}-\frac{1}{\omega_{f i}-\omega+i \epsilon}\right) \\
& =-\frac{i}{\hbar} V_{f i} \frac{2 \epsilon}{\left(\omega_{f i}-\omega\right)^{2}+\epsilon^{2}}
\end{aligned}
$$

We can now take the limit $\epsilon \rightarrow 0$ in which we recover the Dirac $\delta$-function defined in Chapter 4 by the expression (4.6). It gives the final expression for the transition amplitude

$$
\begin{equation*}
a_{f i}=-\frac{i}{\hbar} V_{f i} 2 \pi \delta\left(\omega_{f i}-\omega\right)=-2 \pi i V_{f i} \delta\left(E_{f}-E_{i}-\hbar \omega\right) \tag{6.40}
\end{equation*}
$$

when we make use of $\delta(a x)=(1 / a) \delta(x)$. Now it is clear that this $\delta$-function ensures that energy is conserved in the process in the sense that the final state energy $E_{f}=$ $E_{i}+\hbar \omega$. So we can conclude that the first term in the perturbation (6.38) corresponds to absorption of energy by the system, taken out of the perturbation by the quantum amount $\hbar \omega$. Similarly, the last term in (6.38) will give a transition rate proportional with $\delta\left(E_{f}-E_{i}+\hbar \omega\right)$ which corresponds to emission of energy by the same amount. These relations between the frequencies of the perturbations and the quantized energy differences is just the mathematical proof of Bohr's frequency rules he had to assume in his atomic model for emission and absorption of light.

From the above transition amplitude we can now also calculate the probability for the quantum transition $|i\rangle \rightarrow|f\rangle$ under the perturbation (6.38). Squaring the amplitude, we get a result involving a squared $\delta$-function. But expressiing one of them by the integral (4.7), we can write

$$
P_{f i}=\left|a_{f i}\right|^{2}=\frac{1}{\hbar^{2}}\left|V_{f i}\right|^{2} 2 \pi \delta\left(\omega_{f i}-\omega\right) \int_{-T}^{T} d t e^{\left(\omega_{f i}-\omega\right) t}
$$

again in the limit $T \rightarrow \infty$. But the first $\delta$-function now sets $\omega=\omega_{f i}$ in the integral for the second, thus giving simply the full time $2 T$ the perturbation has lasted, i.e.

$$
P_{f i}=\frac{2 \pi}{\hbar^{2}}\left|V_{f i}\right|^{2} \delta\left(\omega_{f i}-\omega\right) 2 T
$$

Since this probability increases linearly with time, there is a constant number of transitions per unit time. This is expressed by a constant transition rate $\Gamma_{f i}=P_{f i} / 2 T$ which is therefore

$$
\begin{equation*}
\Gamma_{f i}=\frac{2 \pi}{\hbar}\left|V_{f i}\right|^{2} \delta\left(E_{f}-E_{i}-\hbar \omega\right) \tag{6.41}
\end{equation*}
$$

Including higher order terms in the perturbative expansion, this rate will have exactly the same form except for the matrix element $V_{f i}$ being replaced by a more general transition matrix element $T_{f i}=V_{f i}+\cdots$. It is the starting point for many different practical applications and is called Fermi's Golden Rule. From the way it is derived one can understand why it is also valid in relativistic quantum theories.

One might get the impression from that the transition rate is either zero or infinite depending on the frequency $\omega$ having the right value or not compared to the energy difference $E_{f}-E_{i}$. But in practice both the initial and/or the final state will contain a continuum of sub-states in a real situation. Integrating over these, one will have a finite and thus physical result for the transition rate. We will see examples of this in the following.

### 6.7 Rutherford scattering

In a scattering experiment the motion of the free, incoming particles is changed by their mutual interaction. This depends on their relative distance which varies with time. Assuming that the interaction is weak enough, the process can thus be described by time-dependent perturbation theory in quantum mechanics,

When the particles have non-relativistic velocities, we can usually describe this interaction by their mutual, potential energy $V(r)$ where $r$ is their relative separation. Probably the most famous of such scattering experiments were the ones performed by Rutherford and his assistants Marsden and Geiger in 1911. They measured how moving $\alpha$-particles were deflected by heavy atoms like Ag and Au . They observed a surprisingly large number of particles that scattered far away from the incoming direction, This could only be explained by assuming that all the positive charge of the atom was concentrated in a very small nucleus at the center of the atom and holding on to the much lighter electrons moving around at larger distances. This was Rutherford's model for the atom. A few years later it lead N. Bohr to explain the observed spectra of light emission. This was the first triumph of his quantum mechanics.
A schematic illustration of Rutherford's experimental setup is shown in Fig.3.4. The $\alpha$-particles with mass $m$ and velocity $v$ come in from the left towards the atom which we can assume with no loss of generality remains at rest during the process. In the absence of any interactions, each particle would in a classical description pass the nucleus at some shortest distance $b$ shown in the figure. This is the so-called impact parameter


Figure 6.4: Elastic scattering of a classical particle against a potential. The scattering angle $\theta$ is fixed by the impact parameter $b$ and the incoming particle velocity.
related to the angular momentum $L=m v b$ of the particle. Due to the repulsive nuclear Coulomb potential $V(r)=Z e^{2} / 4 \pi \epsilon_{0} r$ acting on the particle at distance $r$, its path is deflected from this straight line by an angle $\theta$ which is the scattering angle. While the momentum $\mathbf{p}$ of the particle now changes both direction and magnitude, the angular momentum $\mathbf{L}=\mathbf{r} \wedge \mathbf{p}$ and the energy $E=\mathbf{p}^{2} / 2 m+V(r)$ remain constant.
In the special case when the impact parameter $b=0$, the particle heads straight towards the nucleus with kinetic energy $E=m v^{2} / 2$. The Coulomb repulsion slows it down and at a distance $R=Z e^{2} / 4 \pi \epsilon_{0} E$ it stops completely up. All its energy is now in the potential part. It then traces the same path back to infinity. The scattering angle is in this very special case $\theta=\pi$.

More interesting is the more general case when the impact parameter is non-zero. The classical trajectory of the particle is then a hyperbola as for unbound Kepler motion. If we denote the direction of the incoming particle seen from the nucleus by the angle $\phi$
shown in the figure, the equation for the trajectory becomes

$$
\begin{equation*}
\frac{1}{r}=\frac{1}{b} \sin \phi-\frac{R}{b^{2}}(1-\cos \phi) \tag{6.42}
\end{equation*}
$$

When the particle has moved away from the nuclues, it continues as a free particle towards the detector at $r \rightarrow \infty$. The corresponding direction is therefore given by $\cos (\phi / 2)=(R / 2 b) \sin (\phi / 2)$ or in terms of the scattering angle $\theta=\pi-\phi$,

$$
\begin{equation*}
\cot \frac{\theta}{2}=\frac{2 b}{R} \tag{6.43}
\end{equation*}
$$

In the limit where $b \rightarrow 0$ it follows that $\theta \rightarrow \pi$ in agreement with what we already have found.

Because of the axial symmetry around the direction of the incoming particles, the distribution of scattered particles will only be a function of the angle $\theta$. Let now $d N(\theta)$ denote the differential number of particles being scattered into the solid angle $d \Omega=2 \pi \sin \theta d \theta$ between the axial directions $\theta$ and $\theta+d \theta$. If $I_{0}$ then is the flux of incoming particles, we have the relation

$$
\begin{equation*}
d N=2 \pi b d b I_{0} \tag{6.44}
\end{equation*}
$$

since these particles must initially have had impact parameters between $b$ and $b+d b$ related to the corresponding scattering angles by (6.43). If we divide this number with the incoming flux, we get a quantity with dimension of area,

$$
\begin{equation*}
d \sigma(\theta)=\frac{d N(\theta)}{I_{0}} \tag{6.45}
\end{equation*}
$$

which is called the differential scattering cross-section. It gives an expression for the number of particles scattered into this particular direction. More explicitly, it can be obtained from

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=\frac{b}{\sin \theta}\left|\frac{d b}{d \theta}\right| \tag{6.46}
\end{equation*}
$$

Now using the relation (6.43) one obtains

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=\left(\frac{Z e^{2}}{16 \pi \epsilon_{0} E}\right)^{2} \frac{1}{\sin ^{4}(\theta / 2)} \tag{6.47}
\end{equation*}
$$

This is Rutherford's famous result for the differential cross-section measured in Coulomb scattering. It is seen to diverge in the forward direction where $\theta=0$. This corresponds to particles with very large or infinite impact parameters. All particles in the incoming beam are therefore scattered and we say that the total or integrated cross section

$$
\begin{equation*}
\sigma=\int_{0}^{4 \pi} d \Omega \frac{d \sigma}{d \Omega} \tag{6.48}
\end{equation*}
$$

is infinite for a pure Coulomb potential. In a real situation for scattering on a neutral atom, the negatively charged electrons will screen off the positive potential from the
atomic nucleus. At large distance the interaction is therefore much weaker and the total scattering cross section is finite. This will be shown in the next section when we consider the same process using quantum mechanics.

The total scattering cross-section gives a quantitative measure for the number of particles scattered out of the incoming beam. If it is along the $x$-axis, the flux $I(x)$ of particles is reduced by an amount $d I=-I \sigma n d x$ by traversing an infinitesemal distance $d x$ of the target with a density of $n$ scattering centers per unit volum. Integrating, this gives a decreasing flux

$$
\begin{equation*}
I(x)=I_{0} e^{-\sigma n x} \tag{6.49}
\end{equation*}
$$

after the beam has traversed a macroscopic distance $x$ of the target. In this way one can measure the total cross-section.

### 6.8 Quantum potential scattering

Let us again consider the situation in Fig, 3.4 where a particle with mass $m$ is scattered by the potential $V(\mathbf{r})$. The full Hamiltonian describing the situation is the usual

$$
\begin{equation*}
\widehat{H}=\frac{1}{2 m} \widehat{\mathbf{p}}^{2}+V(\widehat{\mathbf{r}}) \tag{6.50}
\end{equation*}
$$

where now the potential term induces transitions between the free states described by the first, kinetic part. These are plane waves $\psi=e^{i \mathbf{p} \cdot \mathbf{r} / \hbar}$ with definite momentum $\mathbf{p}$. We can calculate the rate for transitions between these states using the Golden-Rule (6.41). But since now the perturbing potential is static, the frequency $\omega=0$ and there is therefore no change in the total energy of the scattered particle. This is as in classical mechanics.

To be more specific, consider the sitation where the incoming particle has momentum $\mathbf{p}_{i}$. We now want to calculate the rate for transitions to a final state defined by the momentum $\mathbf{p}_{f}$, i.e.

$$
\begin{equation*}
\Gamma_{f i}=\frac{2 \pi}{\hbar}\left|V_{f i}\right|^{2} \delta\left(E_{f}-E_{i}\right) \tag{6.51}
\end{equation*}
$$

Remember that the energies in the $\delta$-function are those in the initial and final states when the potential is zero. It simply ensures energy conservation $E_{i}=E_{f}$ which thus corresponds to $\left|\mathbf{p}_{i}\right|=\left|\mathbf{p}_{f}\right|$. The magnitude of the momentum of the particle is therefore the same in the initial and final states, it is just rotated by the scattering angle $\theta$ as shown in the figure.

Since the apparatus which detects the particle in the final state has a finite opening angle $d \Omega=\sin \theta d \theta d \phi$, it will register all particles with momenta within this range. It corresponds to a density of states

$$
\begin{equation*}
d \rho_{f}=\frac{d^{3} p_{f}}{(2 \pi \hbar)^{3}}=\frac{p_{f}^{2} d p_{f} d \Omega}{(2 \pi \hbar)^{3}} \tag{6.52}
\end{equation*}
$$

and leads to the differential transition rate

$$
d \Gamma=\frac{2 \pi}{\hbar}\left|V_{f i}\right|^{2} \delta\left(E_{f}-E_{i}\right) \frac{p_{f}^{2} d p_{f} d \Omega}{(2 \pi \hbar)^{3}}
$$

into this continuum of states. It is a direct expression for the number of particles that are scattered out of the initial beam. If we divide by the flux $I_{0}$ of incoming particles, we obtain a quantity with dimension of area. This is just the differential cross-section from the definition in the previous section, i.e. $d \sigma=d \Gamma / I_{0}$. Since the final energy $E_{f}^{2}=p_{f}^{2} / 2 m$, it follows that $m d E_{f}=p_{f} d p_{f}$ and we have

$$
d \sigma=\frac{2 \pi}{\hbar I_{0}}\left|V_{f i}\right|^{2} \delta\left(E_{f}-E_{i}\right) \frac{m p_{f} d E_{f}}{(2 \pi \hbar)^{3}} d \Omega
$$

If the incoming particles move along the $x$-axis, the corresponding flux $I_{0}$ is given by

$$
I_{0}=\frac{\hbar}{2 m i}\left(\psi_{i}^{*} \partial_{x} \psi_{i}-\psi_{i} \partial_{x} \psi_{i}^{*}\right)
$$

which for the plane wave $\psi_{i}=e^{i p_{i} x / \hbar}$ simply gives $I_{0}=p_{i} / m$. This is as expected the velocity $v_{i}$ of the particles in their initial state. Finally, we must integrate over the final energy $E_{f}$ since the detector also has a finite energy resolution in addition to its angular resolution. From the $\delta$-function this just isolates those particles with energy $E_{f}=E_{i}$. For the differential cross-section we then have the final result

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=\frac{2 \pi}{\hbar}\left|V_{f i}\right|^{2} \frac{m^{2}}{(2 \pi \hbar)^{3}} \frac{p_{f}}{p_{i}} \tag{6.53}
\end{equation*}
$$

Since the process is elastic, $p_{i}=p_{f} \equiv p$ and cancel out. In addition, for plane waves the matrix element becomes

$$
\begin{equation*}
V_{f i}=\int d^{3} r e^{-i \mathbf{p}_{f} \cdot \mathbf{r} / \hbar} V(\mathbf{r}) e^{i \mathbf{p}_{i} \cdot \mathbf{r} / \hbar}=\int d^{3} r V(\mathbf{r}) e^{-i \mathbf{q} \cdot \mathbf{r} / \hbar} \tag{6.54}
\end{equation*}
$$

where $\mathbf{q}=\mathbf{p}_{f}-\mathbf{p}_{i}$. It is just the 3-dimensional Fourier transform $\tilde{V}(\mathbf{q})$ of the potential. The differential cross-section can thus be written as

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=\left|\frac{m}{2 \pi \hbar^{2}} \tilde{V}(\mathbf{q})\right|^{2} \tag{6.55}
\end{equation*}
$$

When the potential is rotational invariant, i.e. $V=V(r)$, the Fourier transform will only depend on the magnitude $q=|\mathbf{q}|$. The cross-section will then be symmetric around the direction of the incoming particles and thus depend only on the polar scattering angle $\theta$. It is then convenient to introduce the scattering amplitude $f(\theta)=m \tilde{V}(q) / 2 \pi \hbar^{2}$ so that $d \sigma / d \Omega=|f(\theta)|^{2}$. All the angular dependence lies now in the momentum transfer $\mathbf{q}=\mathbf{p}_{f}-\mathbf{p}_{i}$. Taking the square, one obtains $q^{2}=p_{i}^{2}+p_{f}^{2}-2 \mathbf{p}_{f} \cdot \mathbf{p}_{i}=2 p^{2}(1-\cos \theta)$. Thus we have $q=2 p \sin (\theta / 2)$ which also follows geometrically from the kinematics shown in the figure.

In the above we found a flux of incoming particles $I_{0}=v_{i}$ which has usual the dimension of length mer unit time. But a proper flux should have a dimension equal to a number per area per and unit time. This apparent problem arises when we don't insist on using


Figure 6.5: Geometrical determination of the momentum transfer $q$.
correctly normalized states. We should have exactly one particle in the initial state and one in the final. If our system is located within a large, but finite volume $V$, the correct wave functions would be $\psi=e^{i \mathbf{p} \cdot \mathbf{r} / \hbar} / V$. The density of states (6.52) would then be multiplied by $V$ and just be the number of states within the solid angle $d \Omega$ in the final state. Similarly, the flux would then become $I_{0}=v_{i} / V$ which has the correct dimension. But at the same time we see that these two volume factors cancel out when included in the calculation, and our final result for the cross-section remains unchanged.

A physical important example is to consider the screened Coulomb potential

$$
\begin{equation*}
V(r)=\frac{Z e^{2}}{4 \pi \epsilon_{0} r} e^{-\kappa r / \hbar} \tag{6.56}
\end{equation*}
$$

For distance $r \ll R$ where $R=\hbar / \kappa$ is the screening length, it has the standard value while for larger distances it is exponentially damped. This has the same radial dependence as the Yukawa potential which arises in nuclear physics due to the interactions of pions between the nucleons in the nucleus. The screening constant is then $\kappa=m_{\pi} c$ where $m_{\pi}$ is the mass of the pion.

In order to find the differential cross-section for scattering on this potential we need the Fourier transform

$$
\tilde{V}(q)=\frac{Z e^{2}}{\epsilon_{0}} \int d^{3} r \frac{e^{-\kappa r / \hbar}}{4 \pi r} e^{-i \mathbf{q} \cdot \mathbf{r} / \hbar}=\frac{Z e^{2}}{2 \epsilon_{0}} \int_{0}^{\infty} d r r e^{-\kappa r / \hbar} \int_{0}^{\pi} d \beta \sin \beta e^{-i q r \cos \beta / \hbar}
$$

Integrating first over the polar angle $\beta$, one obtains

$$
\tilde{V}(q)=\frac{Z e^{2} i \hbar}{2 q \epsilon_{0}} \int_{0}^{\infty} d r\left[e^{-(\kappa+i q) r / \hbar}-e^{-(\kappa-i q) r / \hbar}\right]
$$

Both of these two remaining exponential integrals are now straightforward and give

$$
\begin{equation*}
\tilde{V}(q)=\frac{Z e^{2}}{\epsilon_{0}} \frac{\hbar^{2}}{q^{2}+\kappa^{2}} \tag{6.57}
\end{equation*}
$$

when combined. From (6.55) we then immediately have the cross-section for scattering on the screened Coulomb potential

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=\left(\frac{Z e^{2} m}{2 \pi \epsilon_{0}}\right)^{2} \frac{1}{\left(4 p^{2} \sin ^{2}(\theta / 2)+\kappa^{2}\right)^{2}} \tag{6.58}
\end{equation*}
$$

after using $q=2 p \sin (\theta / 2)$. When there is no screening, $\kappa=0$ and we recover the standard Rutherford cross-section (6.47) where the energy $E=p^{2} / 2 m$.
One should notice that the Planck-Dirac constant $\hbar$ cancels out in the above quantum calculation of the Rutherford cross-section. Needless to say, this is necessary if it is to be in agreement with the classical result. But this calculation is only in lowest order of perturbation theory. Would not the higher quantum correstions modify the classical result? In fact, one can calculate all these in the special case of a pure Coulomb potential. And they do indeed depend on the quantum constant $\hbar$. But for very deep reasons all these corrections form a geometric series that add up to a simple phase factor in the scattering amplitude. When squared in the cross-section, this factor disappears and one is again left with exactly the classical result.

The differential cross-section (6.58) is finite in the forward direction $\theta=0$ as long as the screening constant $\kappa$ is non-zero. Since it is symmetric in the azimuthal direction, the differential solid angle $d \Omega=2 \pi \sin \theta d \theta$ and the total cross-section will follow from the integral

$$
\sigma=\left(\frac{Z e^{2} m}{2 \pi \epsilon_{0}}\right)^{2} \int_{0}^{\pi} \frac{2 \pi \sin \theta d \theta}{\left(4 p^{2} \sin ^{2}(\theta / 2)+\kappa^{2}\right)^{2}}
$$

Now using $q^{2}=4 p^{2} \sin ^{2}(\theta / 2)$ as a new integration variable with $d q^{2}=2 p^{2} \sin \theta d \theta$, it simplifies to

$$
\sigma=\left(\frac{Z e^{2} m}{2 \pi \epsilon_{0}}\right)^{2} \frac{\pi}{p^{2}} \int_{0}^{4 p^{2}} \frac{d q^{2}}{\left(q^{2}+\kappa^{2}\right)^{2}}=\left(\frac{Z e^{2} m}{2 \pi \epsilon_{0} \kappa}\right)^{2} \frac{\pi}{p^{2}+\kappa^{2} / 4}
$$

As expected, it diverges in the limit $\kappa \rightarrow 0$ of a pure Coulomb potential.
In the low-energy limit $p \rightarrow 0$, the momentum transfer $q \rightarrow 0$ for a potential with finite range as above. The scattering amplitude $f(\theta)$ then becomes a constant

$$
\begin{equation*}
a=\frac{m}{2 \pi \hbar^{2}} \tilde{V}(0) \tag{6.59}
\end{equation*}
$$

independent of the scattering angle $\theta$. It is called the scattering length. The differential cross-section is then simply $d \sigma / d \Omega=a^{2}$ and is uniform in all directions. In the special case of a screened Coulomb potential, we see that $a=Z e^{2} m / 2 \pi \epsilon_{0} \kappa^{2}$. Generally, it is of the same order as the range of the potential or less.

A physical explanation of this simple behaviour of low-energy scattering on finite-range potentials, is obtained by noticing that the incoming particles will then have a de Broglie wave length $\lambda=\hbar / p \gg a$ and the detailed shape of the potential at short distances is therefore no longer 'seen' by the particles. The only observable property of the scattering potential is one parameter, its scattering length.

## Chapter 7

## Electromagnetic interactions

Electrons bound in atoms are held together by the electrostatic Coulomb potential. But there are also important magnetic effects in atoms which will perturb this picture. From classical electrodynamics we know that a charged particle in motion will generate a magnetic field. This will give rise to new interactions in an atom which we will consider here. But the same motion will also in general cause emission of electromagnetic radiation which can cause transitions between the different stationary states of the atom. In the end we will also quantize the electromagnetic field and see how the quanta emerge as massless particles which are the photons.

### 7.1 Magnetic moments

A particle with mass $m$ and electric $e$ moving in an external magnetic field $\mathbf{B}$ feels the Lorentz force

$$
\begin{equation*}
\mathbf{F}=e \mathbf{v} \wedge \mathbf{B} \tag{7.1}
\end{equation*}
$$

Since it is normal to the velocity $\mathbf{v}$, it will bend the motion of the particle, but not do any work. As a result, the particle will move along a circle with radius determined by $e v B=m v^{2} / r$ and angular velocity $\omega=v / r$ which is the cyclotron frequency $\omega_{c}=e B / m$.

We now consider the same particle, but now constrained to move in a circle by some other means. It will then generate a magnetic moment $\boldsymbol{\mu}$ with direction normal to the circle and magnitude $\mu=I A$. Here $I=e / T$ is the current in the loop where $T$ is the time the particle needs to traverse it. It is related to the radius $r$ of the loop and the particle velocity by $v=2 \pi r / T$. For the magnetic moment we thus find $\mu=(e / T) \pi r^{2}=e r v / 2=(e / 2 m) L$ where $L=r m v$ is the angular momentum of the motion. We thus have

$$
\begin{equation*}
\boldsymbol{\mu}=\frac{e}{2 m} \mathbf{L} \tag{7.2}
\end{equation*}
$$

which can be shown to be more generally valid than this simple derivation is.
Let us next place this current loop in the external field $\mathbf{B}$. The particle will still feel the Lorentz force (7.1). But now the force will twist the loop with a torque $\mathbf{N}=\boldsymbol{\mu} \times \mathbf{B}$.

Thus the angular momentum will change according to the standard, mechanical law

$$
\begin{equation*}
\frac{d \mathbf{L}}{d t}=\boldsymbol{\mu} \times \mathbf{B}=\boldsymbol{\omega}_{L} \times \mathbf{L} \tag{7.3}
\end{equation*}
$$

where

$$
\begin{equation*}
\boldsymbol{\omega}_{L}=-\frac{e}{2 m} \mathbf{B} \tag{7.4}
\end{equation*}
$$

is the Larmor frequency. In magnitude, it is one half of the cyclotron frequency $\omega_{c}$. The magnitude of $\mathbf{L}$ will not change because of this, but its direction will precess around the direction of the magnetic field $\mathbf{B}$ with a rate given by Larmor frequency. This precessional motion is equivalent to the classical interaction energy

$$
\begin{equation*}
H_{B}=-\boldsymbol{\mu} \cdot \mathbf{B} \tag{7.5}
\end{equation*}
$$

which we also expect to find in a corresponding version from a quantum derivation.

### 7.2 Vector potentials and gauge invariance

The electric field is the gradient of the scalar potential, $\mathbf{E}=-\boldsymbol{\nabla} \Phi$. From Maxwell's equations we know that the magnetic field always must satisfy $\boldsymbol{\nabla} \cdot \mathbf{B}=0$. It is therefore natural to write it as a curl of a vector potential, $\mathbf{B}=\boldsymbol{\nabla} \wedge \mathbf{A}$ so that this condition is automatically satisfield. Since it is the electric potential $\Phi$ which appears in the Schrödinger equation for a particle with charge $e$ through its potential energy $V=e \Phi$, we should not be surprised to see also the magnetic vector potential A appear in the quantum description of a charged particle moving in a magnetic field. This is in contrast to classical mechanics where it is the electric $\mathbf{E}$ and magnetic field $\mathbf{B}$ that enter the equation of motion via the full Lorentz force

$$
\begin{equation*}
\mathbf{F}=e(\mathbf{E}+\mathbf{v} \wedge \mathbf{B}) \tag{7.6}
\end{equation*}
$$

In quantum mechanics this is not a fundamental quantity, but a derived one as we now will see.

Since the curl of a gradient is always zero, we can always add a gradient to the vector potential without changing the magnetic field. More formally, if the vector potential A is changed into

$$
\begin{equation*}
\mathbf{A} \rightarrow \mathbf{A}^{\prime}=\mathbf{A}+\boldsymbol{\nabla} \chi \tag{7.7}
\end{equation*}
$$

for a scalar function $\chi=\chi(\mathbf{x})$, the magnetic field stays the same,

$$
\begin{equation*}
\mathbf{B} \rightarrow \mathbf{B}^{\prime}=\mathbf{B}+\nabla \wedge \nabla \chi=\mathbf{B} \tag{7.8}
\end{equation*}
$$

The modification (7.7) of the vector potential is called a local gauge transformation and is seen to be a symmetry in the classical theory. It represents the very first encounter with a new, fundamental principle in modern physics from which a basic understanding of the interactions of all particles in Nature can be derived and has been formulated in the Standard Model of elementary particle physics.

As a simple example of a gauge transformation, consider a constant, magnetic field along the $z$-axis, i.e. represented by the vector $\mathbf{B}=(0,0, B)$. It corresponds to the vector potential

$$
\begin{equation*}
\mathbf{A}=\frac{1}{2} \mathbf{B} \wedge \mathbf{r} \tag{7.9}
\end{equation*}
$$

with components $\mathbf{A}=(-y B, x B, 0) / 2$ which is easily verified. Let us now perform a gauge transformation given by $\chi=-B x y / 2$. It results in the transformed vector potential $\mathbf{A}^{\prime}=(-y B, 0,0)$ which obviously cannot be written on the form (7.9). There are in infinity of different potentials corresponding to a given, magnetic field. When we decide to work with one particular of these, we say that we choose a gauge. One such gauge choice is therefore (7.9) which we will make use of later. It is seen to satisfy

$$
\begin{equation*}
\boldsymbol{\nabla} \cdot \mathbf{A}=0 \tag{7.10}
\end{equation*}
$$

which specifies this particular gauge, called the Coulomb gauge. But even with this condition, the gauge is not completely fixed since the above $\mathbf{A}^{\prime}$ and many other potentials also satisfy it.

We are now in the position to consider the quantum mechanics of a charged particle in a magnetic field. The question is how to modify the free Hamiltonian $H=p^{2} / 2 m$ so that it includes the interaction with the field. There is only one requirement, namely that it should reproduce the Lorentz force (7.1) in the classical limit. The correct answer turns out to be the magnetic Hamiltonian

$$
\begin{equation*}
\widehat{H}=\frac{1}{2 m}(\widehat{\mathbf{p}}-e \widehat{\mathbf{A}})^{2} \tag{7.11}
\end{equation*}
$$

where $\widehat{\mathbf{A}}=\mathbf{A}(\widehat{\mathbf{x}})$. It shows that the coupling to the field is found via the simple replacement $\widehat{\mathbf{p}} \rightarrow \widehat{\mathbf{p}}-e \widehat{\mathbf{A}}$ in the free Hamiltonian. This is often called minimal coupling. We will soon see how this follows from imposing gauge invariance also in the quantum theory.

In external field problems it is usually simplest to work in the coordinate representation where the above Hamiltonian takes the form

$$
H=\frac{1}{2 m} \mathbf{p}^{2}-\frac{e}{2 m}(\mathbf{p} \cdot \mathbf{A}+\mathbf{A} \cdot \mathbf{p})+\frac{e^{2}}{2 m} \mathbf{A}^{2}
$$

with $\mathbf{p}=-i \hbar \boldsymbol{\nabla}$. This operator is now supposed to act on some wavefunction $\psi(\mathbf{x})$ to the right. Then

$$
\begin{aligned}
\mathbf{p} \cdot \mathbf{A} \psi & =-i \hbar \boldsymbol{\nabla} \cdot(\mathbf{A} \psi)=-i \hbar(\boldsymbol{\nabla} \cdot \mathbf{A}) \psi-i \hbar \mathbf{A} \cdot \boldsymbol{\nabla} \psi \\
& =-i \hbar \mathbf{A} \cdot \boldsymbol{\nabla} \psi=\mathbf{A} \cdot \mathbf{p} \psi
\end{aligned}
$$

working in the Coulomb gauge (7.10) where the first term vanishes. Thus $\mathbf{A}$ and $\mathbf{p}$ commute in this gauge, which is one reason why it is particularly useful. In the resulting Hamiltonian

$$
\begin{equation*}
H=\frac{1}{2 m} \mathbf{p}^{2}-\frac{e}{m} \mathbf{A} \cdot \mathbf{p}+\frac{e^{2}}{2 m} \mathbf{A}^{2} \tag{7.12}
\end{equation*}
$$

we now insert the vector potential (7.9) in the middle term with the result

$$
\begin{equation*}
H=\frac{1}{2 m} \mathbf{p}^{2}-\frac{e}{2 m} \mathbf{B} \cdot \mathbf{L}+\frac{e^{2}}{2 m} \mathbf{A}^{2} \tag{7.13}
\end{equation*}
$$

since the triple vector product $\mathbf{B} \wedge \mathbf{r} \cdot \mathbf{p}=\mathbf{B} \cdot \mathbf{r} \wedge \mathbf{p}$. We have thus recovered the analog of the classical coupling (7.5) of the magnetic moment (7.2) to the field. The last term in the Hamiltonian is a genuine quantum term, giving rise to atomic diamagnetism.

When the magnetic field is constant along the $z$-axis and using the gauge where the corresponding vector potential is (7.9), the Hamiltonian takes the form

$$
\begin{equation*}
\widehat{H}=\frac{1}{2 m}\left(\widehat{p}_{x}^{2}+\widehat{p}_{y}^{2}\right)+\frac{e^{2} B^{2}}{8 m}\left(\widehat{x}^{2}+\widehat{y}^{2}\right)-\frac{e B}{2 m} \widehat{L}_{z}+\frac{1}{2 m} \widehat{p}_{z}^{2} \tag{7.14}
\end{equation*}
$$

in terms of abstract operators. The motion in the $z$-direction is therefore free as expected. In the $x y$-plane we have an harmonic oscillator with the frequency $\omega=-e B / 2 m$, i.e. the Larmor frequency. Also there is a contribution to the energy from the orbital motion via the angular momentum operator $L_{z}$. In order to disentangle these two effects, it is convenient to introduce the chiral harmonic ladder operators (5.33) $\widehat{a}_{ \pm}$from Chapter 5. Then we have $\widehat{L}_{z}=\hbar\left(\widehat{a}_{+}^{\dagger} \widehat{a}_{+}-\widehat{a}_{-}^{\dagger} \widehat{a}_{-}\right)$with the result that the Hamiltonian simplifies to

$$
\begin{equation*}
\widehat{H}=\hbar \omega\left(2 \widehat{a}_{+}^{\dagger} \widehat{a}_{+}+1\right) \tag{7.15}
\end{equation*}
$$

when we ignore the motion in the $z$-direction. It only depends on the eigenvalue $n_{+}$ of the number operator $\widehat{a}_{+}^{\dagger} \widehat{a}_{+}$which counts the number of quanta in the up-direction, i.e. along the $z$-direction. Each energy level is thus given by just this single quantum number. But since the number $n_{-}$of 'down' quanta can be arbitrary with no effect on the energy, each such energy level has an infinite degeneracy. It is called a Landau level. There is also a non-zero groundstate energy of $E_{0}=\hbar \omega$ as if the motion was confined to a finite region of space even if the $x y$-plane is infinite. This is related to the classical motion in circles. Since the center of such a circle can be everywhere, we also have an intuitive understanding of the infinite degeneracy.

As an extra check we can also now verify that the precessional motion of the magnetic moment comes out right. In the quantum description we must then use the Heisenberg equation of motion

$$
\frac{d \widehat{L}_{i}}{d t}=\frac{i}{\hbar}\left[\widehat{H}, \widehat{L}_{i}\right]
$$

Here we now need the Hamiltonian (7.5) for a magnetic moment at rest in a field. It follows from $(7.5)$ as $\widehat{H}_{B}=-(e / 2 m) \mathbf{B} \cdot \widehat{\mathbf{L}}$ which is consistent with the same term we derived above from the requirement of gauge invariance. Using Einstein's summation convention, we then find follows

$$
\frac{d \widehat{L}_{i}}{d t}=-\frac{i}{\hbar} \frac{e}{2 m} B_{j}\left[\widehat{L}_{j}, \widehat{L}_{i}\right]=-\frac{e}{2 m} \epsilon_{i j k} B_{j} \widehat{L}_{k}=-\frac{e}{2 m}(\mathbf{B} \wedge \widehat{\mathbf{L}})_{i}
$$

in full agreement with the classical result (7.3).

### 7.3 Gauge transformations in quantum theory

It is natural to expect that the invariance of classical physics under the gauge transformation (7.7) is also seen in the quantum theory. But a first look at the Hamiltonian (7.11) is discouraging since it obviously will change. But then we realize that the Hamiltonian always acts on wavefunctions so there may be a way out. In fact, let us consider the stationary Schrödinger equation

$$
\begin{equation*}
\frac{1}{2 m}(\mathbf{p}-e \mathbf{A})^{2} \psi(\mathbf{x})=E \psi(\mathbf{x}) \tag{7.16}
\end{equation*}
$$

combined with the local phase transformation

$$
\begin{equation*}
\psi(\mathbf{x}) \rightarrow \psi^{\prime}(\mathbf{x})=e^{i \theta(\mathbf{x})} \psi(\mathbf{x}) \tag{7.17}
\end{equation*}
$$

Then we find

$$
\begin{aligned}
(\mathbf{p}-e \mathbf{A}) \psi^{\prime}(\mathbf{x}) & =(-i \hbar \boldsymbol{\nabla}-e \mathbf{A}) e^{i \theta(\mathbf{x})} \psi(\mathbf{x}) \\
& =e^{i \theta(\mathbf{x})}(-i \hbar \boldsymbol{\nabla}+\hbar \boldsymbol{\nabla} \theta-e \mathbf{A}) \psi(\mathbf{x})
\end{aligned}
$$

and therefore also

$$
(\mathbf{p}-e \mathbf{A})^{2} \psi^{\prime}(\mathbf{x})=e^{i \theta(\mathbf{x})}(\mathbf{p}+\hbar \boldsymbol{\nabla} \theta-e \mathbf{A})^{2} \psi(\mathbf{x})
$$

But if we now replace the vector potential $\mathbf{A}$ with the transformed vector potential $\mathbf{A}^{\prime}$, use the gauge transformation (7.7) with the until-now arbitrary phase chosen to be $\theta(\mathbf{x})=(e / \hbar) \chi(\mathbf{x})$, we simply obtain

$$
\begin{equation*}
\left(\mathbf{p}-e \mathbf{A}^{\prime}\right)^{2} \psi^{\prime}(\mathbf{x})=e^{i e \chi(\mathbf{x}) / \hbar}(\mathbf{p}-e \mathbf{A})^{2} \psi(\mathbf{x}) \tag{7.18}
\end{equation*}
$$

The Schrödinger equation (7.16) can therefore be made gauge invariant in the sense that it reproduces itself under the simultaneous transformations

$$
\begin{cases}\mathbf{A}(\mathbf{x}) & \rightarrow \mathbf{A}(\mathbf{x})+\boldsymbol{\nabla} \chi(\mathbf{x})  \tag{7.19}\\ \psi(\mathbf{x}) & \rightarrow e^{i e \chi(\mathbf{x}) / \hbar} \psi(\mathbf{x})\end{cases}
$$

They represent local gauge transformations in the quantum theory. Any change of the vector potential of the form (7.7) can be compensated by a simultaneous positiondependent phase transformation of the wave function.

Gauge invariance is actually even deeper. So far we have assumed that the gauge function $\chi$ is independent of time so that the transformed magnetic field remains constant. Only then is the Hamiltonian (7.11) time-independent so that the stationary Schrödinger equation (7.16) have physical content. In the more general case when $\chi=\chi(\mathbf{x}, t)$ we must look for invariance in the time-dependent Schrödinger equation

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t} \Psi(\mathbf{x}, t)=\left[\frac{1}{2 m}(\mathbf{p}-e \mathbf{A})^{2}+e \Phi\right] \Psi(\mathbf{x}, t) \tag{7.20}
\end{equation*}
$$

Here we have also introduced the interaction with an external, electric potential $\Phi=$ $\Phi(\mathbf{x}, t)$ in order to be more general. We now perform the time-dependent gauge transformation

$$
\left\{\begin{align*}
\mathbf{A}(\mathbf{x}, t) & \rightarrow \mathbf{A}(\mathbf{x}, t)+\boldsymbol{\nabla} \chi(\mathbf{x}, t)  \tag{7.21}\\
\Psi(\mathbf{x}, t) & \rightarrow e^{i e \chi(\mathbf{x}, t) / \hbar} \Psi(\mathbf{x}, t)
\end{align*}\right.
$$

Going through the same steps as above, it is then straightforward to show that equation (7.20) is transformed into

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t} \Psi^{\prime}(\mathbf{x}, t)=\left[\frac{1}{2 m}\left(\mathbf{p}-e \mathbf{A}^{\prime}\right)^{2}+e \Phi^{\prime}\right] \Psi^{\prime}(\mathbf{x}, t) \tag{7.22}
\end{equation*}
$$

where now the scalar potential also must undergo a gauge transformation

$$
\begin{equation*}
\Phi(\mathbf{x}, t) \rightarrow \Phi^{\prime}(\mathbf{x}, t)=\Phi(\mathbf{x}, t)-\frac{\partial \chi}{\partial t} \tag{7.23}
\end{equation*}
$$

The last term here is needed in order to cancel a corresponding term on the left-hand side coming from the time derivative of the phase factor in $\Phi^{\prime}$. We thus have established gauge invariance also in the time-dependent Schrödinger theory.

One may think that this scalar transformation will change the electric field given by the gradient of the scalar potential. But in the time-dependent case the electric field is given as

$$
\begin{equation*}
\mathbf{E}=-\nabla \Phi-\frac{\partial \mathbf{A}}{\partial t} \tag{7.24}
\end{equation*}
$$

where the last term is equivalent to Faraday's law. It is also seen to ensure that the electric field remains unchanged under local gauge transformations where now both $\mathbf{A}$ and $\Phi$ will change.

In a covariant description of the electromagnetic field one can combine the transformation of the magnetic vector potential in (7.21) with the transformation of the electric scalar potential in (7.23) into a single transformation of an electromagnetic four-vector potential. That must be left for a later treatment.

### 7.4 The Pauli equation

From Chapter 5 we know that an electron is decribed by a two-component spinor

$$
\begin{equation*}
\psi(\mathbf{x}, t)=\binom{\psi_{\uparrow}(\mathbf{x}, t)}{\psi_{\downarrow}(\mathbf{x}, t)}=\psi_{\uparrow}(\mathbf{x}, t) \alpha+\psi_{\downarrow}(\mathbf{x}, t) \beta \tag{7.25}
\end{equation*}
$$

where $\alpha$ is the eigenspinor for spin along $+z$-axis and $\beta$ for spin in the opposite direction. We can thus interprete $\psi_{\uparrow}(\mathbf{x}, t)$ as the probability amplitude for finding the electron with spin up at position $\mathbf{x}$ at time $t$. Similarly, $\psi_{\downarrow}(\mathbf{x}, t)$ is the corresponding amplitude for finding the opposite spin direction. This position-dependent spinor replaces the ordinary, one-component wave function.

Having only one electron in space corresponds to the normalization

$$
\begin{equation*}
\int d^{3} x \psi^{\dagger} \psi=\int d^{3} x\left(\left|\psi_{\uparrow}\right|^{2}+\left|\psi_{\downarrow}\right|^{2}\right)=1 \tag{7.26}
\end{equation*}
$$

In order to perform these integrals, we must know the space dependence of the spinor $\psi$. It is given by a simple generalization of the Schrödinger equation called the Pauli equation. Recall that the Hamiltonian of a free particle is $H=p^{2} / 2 m$ which we can
write as the diagonal $2 \times 2$ matrix $H=(\boldsymbol{\sigma} \cdot \mathbf{p})^{2} / 2 m$. The spin has no influence on the motion of a free electron.

However, when it moves in an electromagnetic field the situation is more interesting. Requiring also now gauge invariance, we must again make use of minimal coupling $\mathbf{p} \rightarrow \mathbf{p}-e \mathbf{A}$. This results in the Pauli Hamiltonian $H=[\boldsymbol{\sigma} \cdot(\mathbf{p}-e \mathbf{A})]^{2} / 2 m$. Including also an electric potential, we thus have the resulting Schrödinger equation

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t} \psi(\mathbf{x}, t)=\left[\frac{1}{2 m}[\boldsymbol{\sigma} \cdot(\mathbf{p}-e \mathbf{A})]^{2}+e \Phi\right] \psi(\mathbf{x}, t) \tag{7.27}
\end{equation*}
$$

which is called the time-dependent Pauli spinor equation. This Hamiltonian is no longer diagonal, but can be made much more transparent. In order to do that, introduce the vector operator $\boldsymbol{\Pi}=\mathbf{p}-e \mathbf{A}$. It has three components which in general do not commute with each other since $\mathbf{A}=\mathbf{A}(x)$. In fact,

$$
\begin{aligned}
{\left[\Pi_{i}, \Pi_{j}\right] } & =\left[-i \hbar \partial_{i}-e A_{i},-i \hbar \partial_{j}-e A_{j}\right] \\
& =i e \hbar\left(\partial_{i} A_{j}-\partial_{j} A_{i}\right)=i e \hbar F_{i j}
\end{aligned}
$$

which is just curl of the vector potential, i.e. $F_{i j}=\epsilon_{i j k} B_{k}$. We thus have the result $\left[\Pi_{i}, \Pi_{j}\right]=i e \hbar \epsilon_{i j k} B_{k}$ which has the same form as the angular momentum commutator (5.55) in Chapter 5. It can also be written as

$$
\begin{equation*}
\Pi \wedge \Pi=i e \hbar \mathbf{B} \tag{7.28}
\end{equation*}
$$

in the same way as the angular momentum commutator was written as the simple vector product (5.58).

We can now make use of this result in the simplification of the Pauli Hamiltonian (7.27) with help of the general formula (5.82) for the multiplication of sigma-matrices. It gives

$$
\begin{equation*}
(\sigma \cdot \Pi)^{2}=\Pi \cdot \Pi+i \sigma \cdot \Pi \wedge \Pi \tag{7.29}
\end{equation*}
$$

and thus

$$
\begin{equation*}
H=\frac{1}{2 m}(\mathbf{p}-e \mathbf{A})^{2}-\frac{e \hbar}{2 m} \boldsymbol{\sigma} \cdot \mathbf{B}+e \Phi \tag{7.30}
\end{equation*}
$$

for the full Hamiltonian. While the first part is the same as for a spinless particle in (7.11), there is now a new term where the intrinsic spin $\mathbf{S}=(\hbar / 2) \boldsymbol{\sigma}$ couples directly to the magnetic field as a magnetic moment. There is also a magnetic moment coming from the first part due to the orbital motion and derived in (7.13). The total magnetic moment of the electron is therefore

$$
\begin{equation*}
\boldsymbol{\mu}=\frac{e}{2 m}(\mathbf{L}+2 \mathbf{S}) \tag{7.31}
\end{equation*}
$$

The intrinsic part has magnitude $\mu_{e}=e \hbar / 2 m_{e}$ along the $z$-axis. More completely, it can be written as

$$
\begin{equation*}
\boldsymbol{\mu}_{e}=\frac{1}{2} g_{e} \mu_{B} \boldsymbol{\sigma} \tag{7.32}
\end{equation*}
$$

where $\mu_{B}=e \hbar / 2 m_{e}=5.66 \times 10^{-5} \mathrm{eV} / \mathrm{T}$ is a Bohr magneton and $g_{e}=2$ is the electron $g$-factor. It shows that the intrinsic spin contributes twice as much as the orbital part.

This value was historically very important in the establishment of modern quantum mechanics. It is seen here to be a direct consequence of the Pauli Hamiltonian. But at a deeper level it follows from special relativity as formulated in the relativistic Dirac equation for the electron.

More accurate measurements of the magnetic moment of the electron just after the end of WW II showed that its $g$-factor had a slightly different value, $g_{e} / 2=1.00116$. This was explained around the same time by Schwinger who derived the result

$$
\begin{equation*}
g_{e}=2\left(1+\frac{\alpha}{2 \pi}+\cdots\right) \tag{7.33}
\end{equation*}
$$

using quantum electrodynamics, or QED for short. The electromagnetic field is then quantized along with a relativistic spinor field having electrons as quanta. We will say a little more about these theories later.

In a magnetic field the intrinsic spin vector will precess according to the equation of motion (7.3) which is valid for any angular momentum. But since the $g$-factor is close to two, the precessional rate will be twice the Larmor frequency (7.4) and thus very close to the cyclotron frequency $\omega_{c}$. If the electron is in a circular motion with this orbital frequency, the spin vector $\mathbf{S}$ will then rotate with approximately the same angular velocity so that the angle between the momentum and spin is almost constant. Any deviation from the value $g_{e}=2$ will cause this angle to increase with time. This effect allows for a practical measurement of the difference $g_{e}-2$. Nowadays it is put to great use especially for the $\mu$-lepton which is like a heavy, but unstable electron.
One might think that also the proton would be described by the Pauli Hamiltonian (7.30). This is true except for magnetic coupling of the spin. It corresponds to an intrinsic magnetic moment which we can write as

$$
\begin{equation*}
\boldsymbol{\mu}_{p}=\frac{1}{2} g_{p} \mu_{N} \boldsymbol{\sigma} \tag{7.34}
\end{equation*}
$$

where $\mu_{N}=e \hbar / 2 m_{p}$ is a nuclear magneton in analogy with the Bohr magneton. Notice that the charge $e$ is here positive while it was implicitly negative in the corresponding formulas for the electron. But the measured value of this magnetic moment corresponds to a $g$-factor for the proton $g_{p} / 2=2.7928$ which is so different from the ideal value of one that it was a big mystery for many years. This is compounded by the fact that the neutron also has a magnetic moment corresponding to $g_{n} / 2=-1.9135$ even if it has no overall charge.

Today we understand these magnetic moments by the discovery that the nucleons not are fundamental particles, but composed of quarks which are really elementary like the electron. The $u$-quark has charge $2 / 3$ while the $d$-quark has charge $-1 / 3$ in units of the positron charge. In this quark model the proton is composed as $p=u u d$ while the neutron is $n=d d u$. Assuming non-relativistic motion of the quarks and no angular momentum, one then finds a magnetic moment for the proton $\mu_{p}=e \hbar / 2 m_{q}$ and for the neutron $\mu_{n}=-e \hbar / 3 m_{q}$ where $m_{q}$ is the mass of a quark. In this non-relativistic model it is simply given by $m_{p}=m_{n}=3 m_{q}$. Thus we obtain $\mu_{p}=3 \mu_{N}$ and $\mu_{n}=-2 \mu_{N}$ which are within $5 \%$ of the measured values. Why this crude model gives such good results is still not understood.

### 7.5 Zeeman effect

An electron bound in an H -atom which is placed in a magnetic field $\mathbf{B}$, will be perturbed by the dipole interaction (7.5) where the magnetic moment is given by (7.31). If the field is along the $z$-axis, the perturbation becomes

$$
\begin{equation*}
\widehat{H}_{B}=\frac{e B}{2 m_{e}}\left(\widehat{L}_{z}+2 \widehat{S}_{z}\right) \tag{7.35}
\end{equation*}
$$

when we let $e \rightarrow-e$ for an electron with negative charge. The unperturbed states where found in Chapter 5. They are given by the eigenstates $\left|n l m_{l}\right\rangle$ of the Coulomb Hamiltonian and the spin eigenstates $\left|s m_{s}\right\rangle$ of the electron spin with $s=1 / 2$. Combined they form the basis states $\left|n l m_{l}\right\rangle\left|s m_{s}\right\rangle$ which we write more compactly as $\left|n l m_{l} ; m_{s}\right\rangle$. They are eigenstates of $\widehat{L}_{z}$ and $\widehat{S}_{z}$ with respective eigenvalues $\hbar m_{l}$ and $\hbar m_{s}$ where $m_{s}=$ $\pm 1 / 2$. Since the perturbation is diagonal in this basis, we don't have to worry about the degeneracy among the unperturbed states as explained in the previous chapter. In lowest order perturbation theory, each level is then shifted by the amount

$$
\begin{equation*}
E_{B}=\frac{e}{2 m_{e}}\left\langle n l m_{l} m_{s}\right| \widehat{L}_{z}+2 \widehat{S}_{z}\left|n l m_{l} m_{s}\right\rangle=\mu_{B} B\left(m_{l}+2 m_{s}\right) \tag{7.36}
\end{equation*}
$$

which is called the Zeeman effect. The effect is simplest for the ground state 1s. Since it


Figure 7.1: Zeeman splitting of the ground state in hydrogen.
has $\ell=0$, it is simply split into two, depending on the spin quantum number $m_{s}= \pm 1 / 2$ and shown in Fig. 7.2. The difference between the two perturbed levels is $\Delta E_{B}=2 \mu_{B} B$ which is of the order $10^{-4} \mathrm{eV}$ for an external field of magnitude $B=1 \mathrm{~T}$. For the 2 p state the splitting is a bit more complex as shown in in the same figure. Here the orbital quantum number $m_{l}$ will also contribute. The original level is split into four leaving the original level doubly degenerate.

From a physical point of view the effect should be independent of the direction of the magnetic field. Let us illustrate this for the 1 s state with the field along the $x$-axis. Neglecting the orbital part, the Hamiltonian is then simply $H_{B}=\mu_{B} B \sigma_{x}$. It is therefore not diagonal in the standard spin-up $\alpha$ and spin-down $\beta$ basis from Chapter 5 which we have chosen here, i.e.

$$
H_{B}=\mu_{B} B\left(\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right)
$$

We must therefore use degenerate perturbation theory as explained in the previous chapter. The shifts of the unperturbed levels are then given by the eigenvalues of this


Figure 7.2: Zeeman splitting of the 2 p state in hydrogen. Perturbed levels are labelled by ( $m_{l}, m_{s}$ ).
matrix which are easily found to be $\Delta E= \pm \mu_{B} B$. So the spliting of the 1 s level is unchanged as it should be. But the corresponding eigenspinors $\alpha \pm \beta$ now represent states where the spin is along and opposite to the $x$-axis, i.e. parallel and anti-parallel to the magnetic field. This situation is therefore exactly the same as when the field was along the $z$-axis.

Originally one tried to explain these magnetic level shifts by classical physics, but it soon emerged that the effect was much more subtle. A full explanantion was first possible with the advent of quantum mechanics. In fact, the above formula is only correct for very strong magnetic fields $B \gg 1 \mathrm{~T}$. In the next chapter we will see that the motion of the electron in the atom sets up a magnetic field of magnitude $B_{i} \approx 1 \mathrm{~T}$ giving rise to the so-called $L S$-splitting of the levels. It is only when the external field overwhelms this internal field that the formula applies.

### 7.6 Absorption and emission of light

When the external field that perturbs the atom is no longer static, but depends on time, we have an even more important system corresponding to atoms absorbing and emitting electromagnetic radiation. So far we don't have a quantum theory of light so let us first consider electromagnetic radiation which can be described by classical theory. In certain limits this should correspond to a full quantum treatment we will come back to later.
Free, electromagnetic radiation is described by the vector potential $\mathbf{A}(\mathbf{x}, t)$ alone. The electric $\mathbf{E}$ and magnetic fields $\mathbf{B}$ will then be given by

$$
\begin{equation*}
\mathbf{E}(\mathbf{x}, t)=-\frac{\partial}{\partial t} \mathbf{A}(\mathbf{x}, t), \quad \mathbf{B}(\mathbf{x}, t)=\boldsymbol{\nabla} \wedge \mathbf{A}(\mathbf{x}, t) \tag{7.37}
\end{equation*}
$$

A homogeneous radiation field will be made up of unpolarized waves in all directions. So let us first consider one of these waves with a definite wave number $\mathbf{k}$ and polarization vector $\mathbf{e}_{\lambda}$ so that

$$
\begin{equation*}
\mathbf{A}(\mathbf{x}, t)=a \mathbf{e}_{\lambda} \cos (\mathbf{k} \cdot \mathbf{x}-\omega t) \tag{7.38}
\end{equation*}
$$

where $a$ is the amplitude and $\omega=c|\mathbf{k}|$. The Coulomb gauge condition $\boldsymbol{\nabla} \cdot \mathbf{A}=0$ now implies that $\mathbf{k} \cdot \mathbf{e}_{\lambda}=0$ so that the polarization vector is normal to the wavevector. There are two such independent polarization vectors and it is convenient to take them


Figure 7.3: Polarization vectors for wave in direction of the wave number $\mathbf{k}$ can be along the electric and magnetic fields, but always normal to the wavevector.
to be orthonormal to each other, i.e. $\mathbf{e}_{\lambda} \cdot \mathbf{e}_{\lambda^{\prime}}=\delta_{\lambda \lambda^{\prime}}$. Notice that the directions of the polarization vectors depend intimately on the direction of the wave number. For the electric field we now find

$$
\mathbf{E}=-a \omega \mathbf{e}_{\lambda} \sin (\mathbf{k} \cdot \mathbf{x}-\omega t)
$$

It is along one of the polarization vector while the magnetic field

$$
\mathbf{B}=-a \mathbf{k} \wedge \mathbf{e}_{\lambda} \sin (\mathbf{k} \cdot \mathbf{x}-\omega t)
$$

is normal to it, i.e. along the other one.
In general the energy density in the radiation is

$$
\begin{equation*}
u=\frac{1}{2}\left\langle\epsilon_{0} \mathbf{E}^{2}+\mathbf{B}^{2} / \mu_{0}\right\rangle \tag{7.39}
\end{equation*}
$$

where the average is over time and space. Since we are considering a single wave with a definite frequency $\omega$, this will in our case be a spectral density. Using now that $\left\langle\sin ^{2}\right\rangle=\left\langle\cos ^{2}\right\rangle=1 / 2$, we get $\left\langle\mathbf{E}^{2}\right\rangle=a^{2} \omega^{2} / 2$ and $\left\langle\mathbf{B}^{2}\right\rangle=a^{2} k^{2} / 2$ and therefore

$$
u=\frac{1}{2} \cdot \frac{a^{2}}{2}\left(\epsilon_{0} \omega^{2}+\frac{1}{\mu_{0}} k^{2}\right)=\frac{1}{2} \epsilon_{0} a^{2} \omega^{2}
$$

since $c^{2}=1 / \epsilon_{0} \mu_{0}$. We thus find for the amplitude $a=\sqrt{2 u / \epsilon_{0} \omega^{2}}$ when expressed in terms of the spectral energy density.
For simplicity we again consider an atom with a single electron bound to a nuclear charge $Z e$ in the presence of such electromagnetic radiation. Ignoring here the spin of the elctron, the corresponding Hamiltonian is

$$
\begin{equation*}
H=\frac{1}{2 m_{e}}(\mathbf{p}+e \mathbf{A})^{2}-\frac{Z e^{2}}{4 \pi \epsilon_{0} r} \tag{7.40}
\end{equation*}
$$

It can be written as $H=H_{0}+V$ with the unperturbed part

$$
\begin{equation*}
H_{0}=\frac{1}{2 m_{e}} \mathbf{p}^{2}-\frac{Z e^{2}}{4 \pi \epsilon_{0} r} \tag{7.41}
\end{equation*}
$$

with eigenstates $|n\rangle$ and corresponding energies $E_{n}$ plus the perturbing part

$$
\begin{equation*}
V=\frac{e}{m_{e}} \mathbf{A} \cdot \mathbf{p}+\frac{e^{2}}{2 m_{e}} \mathbf{A}^{2} \tag{7.42}
\end{equation*}
$$

We are again in the the Coulomb gauge where $\mathbf{A} \cdot \mathbf{p}=\mathbf{p} \cdot \mathbf{A}$. The perturbing part is driven by the vector potential $\mathbf{A}$ which we will consider in first order perturbation theory. It is therefore consistent to neglect the term $\propto \mathbf{A}^{2}$ which would give second order effects. Writing the vecor potential (7.38) as

$$
\begin{equation*}
\mathbf{A}(\mathbf{x}, t)=\frac{1}{2} a \mathbf{e}_{\lambda}\left[e^{i(\mathbf{k} \cdot \mathbf{x}-\omega t)}+e^{i(\mathbf{k} \cdot \mathbf{x}-\omega t)}\right] \tag{7.43}
\end{equation*}
$$

we know from the previous chapter on time-dependent perturbation theory that the first term will induce transitions in the atom corresponding to absorption of an energy amount $\hbar \omega$, while the second term corresponds to emission of the same energy amount. To be specific, assume the atom is initially in a state $|n\rangle$ and we want to find the


Figure 7.4: When radiation is absorbed, the atom makes a transition from a lower state to a higher state.
transition rate to a higher energy state $|m\rangle$ as illustrated in the Fig.7.4. The absorption rate is then given by Fermi's Golden Rule as

$$
\begin{equation*}
\Gamma_{m n}=\frac{2 \pi}{\hbar}\left|V_{m n}\right|^{2} \delta\left(E_{m}-E_{n}-\hbar \omega\right) \tag{7.44}
\end{equation*}
$$

where the transition matrix element is

$$
V_{m n}=\frac{e}{m_{e}} a\langle m| e^{i \mathbf{k} \cdot \widehat{\mathbf{x}}} \mathbf{e}_{\lambda} \cdot \widehat{\mathbf{p}}|n\rangle
$$

Now we have $\left\langle e^{i \mathbf{k} \cdot \mathbf{x}}\right\rangle=1+i\langle\mathbf{k} \cdot \mathbf{x}\rangle+\ldots$ where $\langle\mathbf{k} \cdot \mathbf{x}\rangle \leq\langle k\rangle\langle r\rangle$. Since $\langle r\rangle$ is of the order of the Bohr radius $a=Z / \alpha m_{e} c$ and $k=\omega / c$, we see that $\langle\mathbf{k} \cdot \mathbf{x}\rangle \approx(Z \alpha)^{3} \ll 1$. Thus we can ignore the exponential term. We have then made the dipole approximation which is seen to be very accurate here. The simplified transition element

$$
V_{m n}=\frac{e}{m_{e}} a \mathbf{e}_{\lambda} \cdot\langle m| \hat{\mathbf{p}}|n\rangle
$$

we can now either calculate directly from the known hydrogen eigenfunctions or make use of the Heisenberg equation of motion $\widehat{\mathbf{p}}=i m_{e}\left[\widehat{H}_{0}, \widehat{\mathbf{x}}\right]$. It allows us to write the matrix element

$$
\begin{align*}
\langle m| \widehat{\mathbf{p}}|n\rangle & =i m_{e}\langle m| \widehat{H}_{0} \widehat{\mathbf{x}}-\widehat{\mathbf{x}} \widehat{H}_{0}|n\rangle / \hbar  \tag{7.45}\\
& =i m_{e}\left(E_{m}-E_{n}\right)\langle m| \widehat{\mathbf{x}}|n\rangle \hbar=i m_{e} \omega \mathbf{x}_{m n} \tag{7.46}
\end{align*}
$$

where $\mathbf{x}_{m n}=\langle m| \hat{\mathbf{x}}|n\rangle$. Here we have made use of the $\delta$-function in (7.44) which fixes the frequency $\omega$ in terms of the atomic energy difference $E_{m}-E_{n}$. This is just the Bohr frequency formula. For these two particular initial and final states, only radiation with this frequency will be absorbed.

With these simplifications the absorption rate can now be written as

$$
\Gamma_{m n}=\frac{\pi e^{2}}{2 \hbar}(a \omega)^{2}\left|\mathbf{e}_{\lambda} \cdot \mathbf{x}_{m n}\right|^{2} \delta\left(E_{m}-E_{n}-\hbar \omega\right)
$$

or

$$
\begin{equation*}
\Gamma_{m n}=\frac{\pi e^{2}}{\epsilon_{0} \hbar} u(\omega)\left|\mathbf{e}_{\lambda} \cdot \mathbf{x}_{m n}\right|^{2} \delta\left(E_{m}-E_{n}-\hbar \omega\right) \tag{7.47}
\end{equation*}
$$

when we use our previous result for the wave amplitude $a$ expressed in terms of the spectral energy density $u(\omega)$.

It is now easy to generalize this result for the absorption rate from a single wave to more general radiation. We must then average over all directions of the wavenumber $\mathbf{k}$ and integrate over all frequencies $\omega$. While this integration simply picks out the frequency $\omega=\left(E_{m}-E_{n}\right) / \hbar$, the first average corresponds to an average over all directions of the polarization vector $\mathbf{e}_{\lambda}$. This simply gives

$$
\begin{equation*}
\left.\left.\langle | \mathbf{e}_{\lambda} \cdot \mathbf{x}_{m n}\right|^{2}\right\rangle=\frac{1}{3}\left|\mathbf{x}_{m n}\right|^{2} \tag{7.48}
\end{equation*}
$$

since all three directions contribute equally for isotropic radiation. For a more elaborate, mathematical proof, choose coordinate axes so that the wavenumber $\mathbf{k}$ is along the $z$ axis and the vector $\left|\mathbf{x}_{m n}\right|$ in the $y z$-plane at an angle $\theta$ away from the $z$-axis.. The polarization vector $\mathbf{e}_{\lambda}$ is then in the $x y$-plane forming an angle $\phi$ with the $x$-axis. Then $\left|\mathbf{e}_{\lambda} \cdot \mathbf{x}_{m n}\right|=\left|\mathbf{x}_{m n}\right| \sin \theta \sin \phi$. Averaging over all directions of $\mathbf{k}$, now corresponds to integrating these spherical angles over the unit sphere, i.e.

$$
\begin{aligned}
\left.\left.\langle | \mathbf{e}_{\lambda} \cdot \mathbf{x}_{m n}\right|^{2}\right\rangle & =\frac{1}{4 \pi} \int_{0}^{\pi} d \theta \sin \theta \int_{0}^{2 \pi} d \phi\left(\left|\mathbf{x}_{m n}\right| \sin \theta \sin \phi\right)^{2} \\
& =\frac{1}{4 \pi}\left|\mathbf{x}_{m n}\right|^{2} \int_{0}^{\pi} d \theta \sin ^{3} \theta \int_{0}^{2 \pi} d \phi \sin ^{2} \phi \\
& =\frac{1}{4 \pi}\left|\mathbf{x}_{m n}\right|^{2} \cdot \frac{4}{3} \cdot \pi=\frac{1}{3}\left|\mathbf{x}_{m n}\right|^{2}
\end{aligned}
$$

Introducing again the fine-structure constant $\alpha=e^{2} / 4 \pi \epsilon_{0} \hbar c$, the resulting rate valid for isotropic radiation is then

$$
\begin{equation*}
\Gamma_{m n}=\frac{4 \pi^{2} \alpha c}{3 \hbar}\left|\mathbf{x}_{m n}\right|^{2} u(\omega) \tag{7.49}
\end{equation*}
$$

where again it should be remembered that the frequency is fixed by the above Bohr condition.

Emission of radiation will be caused by the second term of the vector potential in (7.43). The atom is then initially in the upper state $|m\rangle$ and can go to the lower state $|n\rangle$ at


Figure 7.5: When radiation is emitted, the atom makes a transition from an upper state to a lower state.
the rate

$$
\begin{equation*}
\Gamma_{n m}=\frac{2 \pi}{\hbar}\left|V_{n m}\right|^{2} \delta\left(E_{n}-E_{m}+\hbar \omega\right) \tag{7.50}
\end{equation*}
$$

where now the transition matrix element is

$$
\begin{equation*}
V_{n m}=\frac{e}{m_{e}} a\langle n| e^{-i \mathbf{k} \cdot \widehat{\mathbf{x}}} \mathbf{e}_{\lambda} \cdot \widehat{\mathbf{p}}|m\rangle \tag{7.51}
\end{equation*}
$$

Again making use of the dipole approximation and the Heisenberg equation of motion, it gives

$$
\begin{equation*}
\Gamma_{n m}=\frac{4 \pi^{2} \alpha c}{3 \hbar}\left|\mathbf{x}_{n m}\right|^{2} u(\omega) \tag{7.52}
\end{equation*}
$$

for the emission rate after averaging over polarizations. But now we see that this equals the absorption rate (7.49). This follows from $\left|\mathbf{x}_{n m}\right|^{2}=\langle n| \widehat{\mathbf{x}}|m\rangle\langle n| \widehat{\mathbf{x}}|m\rangle^{*}$. Now using that the operator $\widehat{\mathbf{x}}$ is Hermitean, we have $\left|\mathbf{x}_{n m}\right|^{2}=\langle m| \widehat{\mathbf{x}}|n\rangle\langle m| \widehat{\mathbf{x}}|n\rangle^{*}=\left|\mathbf{x}_{m n}\right|^{2}$ and therefore

$$
\begin{equation*}
\Gamma_{m n}=\Gamma_{n m} \tag{7.53}
\end{equation*}
$$

This important result is called detailed balance. It can be generalized to other quantum transitions and has important implications in subatomic physics and statistical mechanics.

### 7.7 Einstein A and B coefficients

That the ratse for emission and absorption of electromagnetic radiation are equal, expresses a form of equilibrium. This was also on Einstein's mind in 1917 when he contemplated the processes going on with atoms in thermal equlibrium with blackbody radiation with temperature $T$. According to the atomic model of Bohr from 1913, he
knew that the atoms could be stable in stationary states with quantized energies. Considering the numbers $N_{m}$ and $N_{n}$ of atoms with energies $E_{m}$ and $E_{n}$, these had to be related by the Boltzmann distribution

$$
\begin{equation*}
\frac{N_{m}}{N_{n}}=e^{-\beta\left(E_{m}-E_{n}\right)} \tag{7.54}
\end{equation*}
$$

where $\beta=1 / k_{B}$ and $k_{B}$ is the Boltzmann constant. This ratio had to remain constant even if there constantly were transitions between these two states. He furthermore assumed that the transition rate for absorption could be written as $\Gamma_{m n}=C u(\omega)$ where the factor $C$ should be independent of temperature while $u(\omega)$ was the blackbody spectral energy density

$$
\begin{equation*}
u(\omega)=\frac{\omega^{2}}{\pi^{2} c^{3}} \frac{\hbar \omega}{e^{\beta \hbar \omega}-1} \tag{7.55}
\end{equation*}
$$

found by Planck in 1900. The total rate for upward transitions $n \rightarrow m$ is therefore $\Gamma(n \rightarrow m)=N_{n} C u$. In thermal equlibrium this must equal the total downward rate $\Gamma(m \rightarrow n)=N_{m} \Gamma_{n m}$. But now he discovered that the equilibrium condition

$$
\begin{equation*}
\Gamma(n \rightarrow m)=\Gamma(m \rightarrow n) \tag{7.56}
\end{equation*}
$$

can only be satisfied when the emission rate has the form $\Gamma_{n m}=A+B u(\omega)$ where the coefficients $A$ and $B$ again are independent of temperature and determined purely by atomic physics. Since the transitions caused by $A$ is independent of the radiation, it is


Figure 7.6: When radiation is emitted, the atom makes a transition from an upper state to a lower state.
said to describe the rate for spontaneous emission. The coefficient $B$, on the other hand, induces downward transitions because of radiation already being present and is therefore said to be the rate for induced emission. It corresponds to the rate we calculated in the previous section. These different transitions and accompanying coefficients lay at the heart of modern laser physics and can be illustrated as in the above figure.

We can now write the equilibrium condition as

$$
\frac{N_{n}}{N_{m}}=\frac{A+B u}{C u}=\frac{A}{C u}+\frac{B}{C}
$$

Since $\hbar \omega=E_{m}-E_{n}$, we can write the left-hand side as $\exp (\beta \hbar \omega)$ and therefore

$$
e^{\beta \hbar \omega}=\frac{A}{C} \frac{\pi^{2} c^{3}}{\omega^{2}} \frac{e^{\beta \hbar \omega}-1}{\hbar \omega}+\frac{B}{C}
$$

Needless to say, this must be true at all temperatures. The exponential term on the left must therefore match the expontial term on the right which implies the first result $A \pi^{2} c^{3}=C \hbar \omega^{3}$ or

$$
\begin{equation*}
A=\frac{\hbar \omega^{3}}{\pi^{2} c^{3}} C \tag{7.57}
\end{equation*}
$$

The remaining two terms on the right must cancel and therefore give $B=C$. This just expresses the same result of detailed balance (7.53) as we derived in in the previous section from the machinery of quantum mechanics. Remember that Einstein in 1917 didn't know anything of that sort.

But he also had the new result (7.57) which gives the rate for emission from an atom in the absence of any other initial radiation. If we now extract the coefficient $C$ from the absorption rate (7.49), we thus find for the spontaneous emission rate the important result

$$
\begin{equation*}
\Gamma_{n m}^{\text {spont }}=A=\frac{4 \alpha \omega^{3}}{3 c^{2}}\left|\mathbf{x}_{m n}\right|^{2} \tag{7.58}
\end{equation*}
$$

In the following section we will explicitly calculate this rate for the lowest transition in the hydrogen atom which is the simplest case.

A general excited state can make such radiative transitions to many lower states. We call the sum over all these partial transition rates for the total decay rate $\Gamma$. If the number of atoms in this excited state is $N$, a number $d N=-N \Gamma d t$ will decay in an infinitesemal time interval $d t$. Integrating, this gives the exponential decay law

$$
\begin{equation*}
N=N_{0} e^{-\Gamma t} \tag{7.59}
\end{equation*}
$$

where $N_{0}$ is the initial number at time $t=0$. After a time $\tau=1 / \Gamma$ called the lifetime of this state, the number has fallen to $N_{0} / e$. It is easy to make a rough estimate of this lifetime in hydrogen with $Z=1$. Since the matrix element $\left|\mathbf{x}_{m n}\right| \approx a_{0}=1 / \alpha m_{e} c$ and $\omega \approx m_{e} c^{2} \alpha^{2} / \hbar$, it follows that $\Gamma \approx \alpha^{5} m_{e} c^{2} / \hbar \approx 10^{-7} \mathrm{eV} / \hbar$ which gives $\tau \approx 10^{-10} \mathrm{sec}$. The lifetime is very short, but still long to what we are used to in the macroscopic world.

Before we present a more accurate calculation, we will give a quantum description of the electromagnetic field itself. It will give us a more fundamental understanding of why these spontaneous transitions take place after all. This is at the core of the quantum mystery.

### 7.8 Quantization of the electromagnetic field

It simplest to consider the vector field $\mathbf{A}(\mathbf{x}, t)$ within a cubic box of volume $V=L^{3}$ and satisfying periodic boundary conditions. The field can then be expanded in a Fourier series

$$
\begin{equation*}
\mathbf{A}(\mathbf{x}, t)=\sqrt{\frac{1}{V}} \sum_{\mathbf{k}} \mathbf{A}_{\mathbf{k}}(t) e^{i \mathbf{k} \cdot \mathbf{x}} \tag{7.60}
\end{equation*}
$$

where the allowed wavenumbers can be written as $\mathbf{k}=2 \pi \mathbf{n} / L$ where the direction vector $\mathbf{n}=\left(n_{x}, n_{y}, n_{z}\right)$ with $n_{i}=0, \pm 1, \pm 2, \ldots$. To verify this, let for example one coordinate $x$ be increased to $x \rightarrow x+L$. The mode function $\exp (i k x)$ will then change to $\exp (i k x) \rightarrow \exp (i k x+i k L)=\exp (i k x)$ since $\exp (i k L)=\exp (2 \pi i n)=1$. These mode functions are also orthonormalized as seen from

$$
\int_{0}^{L} d x e^{i\left(k-k^{\prime}\right) x}=\left.\frac{e^{i\left(k-k^{\prime}\right) x}}{i\left(k-k^{\prime}\right)}\right|_{0} ^{L}=\frac{e^{i\left(k-k^{\prime}\right) L}-1}{i\left(k-k^{\prime}\right)}= \begin{cases}L, & k=k^{\prime}  \tag{7.61}\\ 0, & k \neq k^{\prime}\end{cases}
$$

since $k-k^{\prime}=2 \pi\left(n-n^{\prime}\right) / L$.
In the limit when the volume $V=L^{3}$ becomes very large, we can replace the discrete sum over modes my a corresponding integral, noticing that $\sum_{\mathbf{k}}=\sum_{\mathbf{n}} \Delta n_{x} \Delta n_{y} \Delta n_{z}$ with $\Delta n_{i}=1$. We thus have

$$
\begin{equation*}
\sum_{\mathbf{k}}=\left(\frac{L}{2 \pi}\right)^{3} \sum_{\mathbf{k}} \Delta k_{x} \Delta k_{y} \Delta k_{z} \rightarrow V \int \frac{d^{3} k}{(2 \pi)^{3}} \tag{7.62}
\end{equation*}
$$

This is a very useful formula in statistical physics and quantum mechanics.
The Fourier modes $\mathbf{A}_{\mathbf{k}}$ in (7.60) are in general complex, but satisfy simple reality conditions. This follows from the requirement that the field $\mathbf{A}(\mathbf{x}, t)$ is real so that

$$
\mathbf{A}^{*}(\mathbf{x}, t)=\sqrt{\frac{1}{V}} \sum_{\mathbf{k}} \mathbf{A}_{\mathbf{k}}^{*}(t) e^{-i \mathbf{k} \cdot \mathbf{x}}=\mathbf{A}(\mathbf{x}, t)
$$

Making here the replacement $\mathbf{k} \rightarrow-\mathbf{k}$ in the sum, it follows that $\mathbf{A}_{\mathbf{k}}^{*}=\mathbf{A}_{-\mathbf{k}}$. We will often make use of it in the following.

The total energy in the field is now

$$
\begin{equation*}
E=\int d^{3} x\left(\frac{1}{2} \epsilon_{0} \mathbf{E}^{2}+\frac{1}{2 \mu_{0}} \mathbf{B}^{2}\right) \tag{7.63}
\end{equation*}
$$

where the electric and magnetic fields are given by (7.37). For the electric field we find

$$
\mathbf{E}(\mathbf{x}, t)=-\frac{\partial}{\partial t} \mathbf{A}(\mathbf{x}, t)=-\sqrt{\frac{1}{V}} \sum_{\mathbf{k}} \dot{\mathbf{A}}_{\mathbf{k}}(t) e^{i \mathbf{k} \cdot \mathbf{x}}
$$

and therefore

$$
\int d^{3} x \mathbf{E}^{2}=\frac{1}{V} \sum_{\mathbf{k}, \mathbf{k}^{\prime}} \dot{\mathbf{A}}_{\mathbf{k}} \dot{\mathbf{A}}_{\mathbf{k}^{\prime}} \int d^{3} x e^{i \mathbf{k} \cdot \mathbf{x}} e^{i \mathbf{k}^{\prime} \cdot \mathbf{x}}
$$

But now we make use of the periodic boundary conditions satisfied by the exponential mode functions. Indeed, the generalized version of (7.61) is now

$$
\begin{equation*}
\int d^{3} x e^{i\left(\mathbf{k}-\mathbf{k}^{\prime}\right) \cdot \mathbf{x}}=V \delta_{\mathbf{k}, \mathbf{k}^{\prime}} \tag{7.64}
\end{equation*}
$$

which now allows us to obtain

$$
\int d^{3} x \mathbf{E}^{2}=\sum_{\mathbf{k}} \dot{\mathbf{A}}_{\mathbf{k}} \dot{\mathbf{A}}_{-\mathbf{k}}=\sum_{\mathbf{k}} \dot{\mathbf{A}}_{\mathbf{k}} \dot{\mathbf{A}}_{\mathbf{k}}^{*}=\sum_{\mathbf{k}}\left|\dot{\mathbf{A}}_{\mathbf{k}}\right|^{2}
$$

Similarly, for the magnetic field

$$
\mathbf{B}(\mathbf{x}, t)=\boldsymbol{\nabla} \wedge \mathbf{A}(\mathbf{x}, t)=\sqrt{\frac{1}{V}} \sum_{\mathbf{k}} i \mathbf{k} \wedge \mathbf{A}_{\mathbf{k}}(t) e^{i \mathbf{k} \cdot \mathbf{x}}
$$

it follows that

$$
\int d^{3} x \mathbf{E}^{2}=\sum_{\mathbf{k}}\left(i \mathbf{k} \wedge \mathbf{A}_{\mathbf{k}}\right) \cdot\left(-i \mathbf{k} \wedge \mathbf{A}_{-\mathbf{k}}\right)
$$

Using now the vector identity $(\mathbf{A} \wedge \mathbf{B}) \cdot(\mathbf{C} \wedge \mathbf{D})=(\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D})$ combined with the Coulomb gauge requirement $\boldsymbol{\nabla} \cdot \mathbf{A}=0$ which gives $\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}}=0$, it then follows that $\left(\mathbf{k} \wedge \mathbf{A}_{\mathbf{k}}\right) \cdot\left(\mathbf{k} \wedge \mathbf{A}_{-\mathbf{k}}\right)=\mathbf{k}^{2} \mathbf{A}_{\mathbf{k}} \cdot \mathbf{A}_{-\mathbf{k}}$ which then gives the magnetic energy. The total energy (7.63) then becomes

$$
\begin{equation*}
E=\frac{\epsilon_{0}}{2} \sum_{\mathbf{k}}\left(\dot{\mathbf{A}}_{\mathbf{k}} \dot{\mathbf{A}}_{\mathbf{k}}^{*}+\omega_{\mathbf{k}}^{2} \mathbf{A}_{\mathbf{k}} \cdot \mathbf{A}_{\mathbf{k}}^{*}\right) \tag{7.65}
\end{equation*}
$$

when we again make use of $c^{2}=1 / \epsilon_{0} \mu_{0}$. Here $\omega_{\mathbf{k}}=c|\mathbf{k}|$ is the frequency associated with the wavevector $\mathbf{k}$.

We have already made use of the fact that the field mode $\mathbf{A}_{\mathbf{k}}$ is orthogonal to the wavevector $\mathbf{k}$. That is always the case in the Coulomb gauge. The vector $\mathbf{A}_{\mathbf{k}}$ is therefore situated in a plane normal to $\mathbf{k}$. It is then convenient to introduce two orthogonal, unit vectors $\mathbf{e}_{\lambda}=\mathbf{e}_{\lambda}(\mathbf{k})$ in this plane. They can in general be complex and thus satisfies

$$
\begin{equation*}
\mathbf{e}_{\lambda}^{*} \cdot \mathbf{e}_{\lambda^{\prime}}=\delta_{\lambda \lambda^{\prime}} \tag{7.66}
\end{equation*}
$$

Needless to say, these are the polarization vectors already introduced earlier in this chapter. Together with the unit vector $\widehat{\mathbf{k}}$ in the direction of the wavevector $\mathbf{k}$ they form a complete set. Thus we have $\sum_{\lambda} e_{\lambda i} e_{\lambda j}+\widehat{k}_{i} \widehat{k}_{j}=\delta_{i j}$ or

$$
\begin{equation*}
\sum_{\lambda} e_{\lambda i} e_{\lambda j}=\delta_{i j}-k_{i} k_{j} / \mathbf{k}^{2} \tag{7.67}
\end{equation*}
$$

which is a very useful result in practical calculation.
With these polarization vectors we have the decomposition

$$
\mathbf{A}_{\mathbf{k}}(t)=\sum_{\lambda} A_{\mathbf{k} \lambda}(t) \mathbf{e}_{\lambda}
$$

where the field mode $A_{\mathbf{k}, \lambda}$ now is scalar, but still complex. This allows us to rewrite the product

$$
\mathbf{A}_{\mathbf{k}} \cdot \mathbf{A}_{\mathbf{k}}^{*}=\sum_{\lambda \lambda^{\prime}} A_{\mathbf{k} \lambda} \mathbf{e}_{\lambda} \cdot \mathbf{e}_{\lambda^{\prime}}^{*} A_{\mathbf{k} \lambda^{\prime}}^{*}=\sum_{\lambda} A_{\mathbf{k} \lambda} A_{\mathbf{k} \lambda}^{*}
$$

and therefore also the energy as

$$
\begin{equation*}
E=\frac{\epsilon_{0}}{2} \sum_{\mathbf{k} \lambda}\left(\dot{A}_{\mathbf{k} \lambda} \dot{A}_{\mathbf{k} \lambda}^{*}+\omega_{\mathbf{k}}^{2} A_{\mathbf{k} \lambda} A_{\mathbf{k} \lambda}^{*}\right) \tag{7.68}
\end{equation*}
$$

Each mode specified by $(\mathbf{k}, \lambda)$ contributes a term to the energy which has exactly the same form as the energy of a 2-dimensional, harmonic oscillator,

$$
E=\frac{1}{2} m \dot{z} \dot{z}^{*}+\frac{1}{2} m \omega^{2} z z^{*}
$$

where $z=(x+i y) / \sqrt{2}$. We can therefore consider the field component $A_{\mathbf{k} \lambda}(t)$ as a dynamical variable like the oscillator coordinate $z(t)$. The conjugate momentum variable is $\Pi_{\mathbf{k} \lambda}=\epsilon_{0} \dot{A}_{\mathbf{k} \lambda}^{*}$. Quantization of the electromagnetic field is now achieved by letting these two variables become quantum operators,

$$
\begin{equation*}
A_{\mathbf{k} \lambda}(t) \longrightarrow \widehat{A}_{\mathbf{k} \lambda}(t), \quad \Pi_{\mathbf{k} \lambda}(t) \longrightarrow \widehat{\Pi}_{\mathbf{k} \lambda}(t) \tag{7.69}
\end{equation*}
$$

with the canonical commutators

$$
\begin{equation*}
\left[\widehat{A}_{\mathbf{k} \lambda}, \widehat{A}_{\mathbf{k}^{\prime} \lambda^{\prime}}\right]=0, \quad\left[\widehat{A}_{\mathbf{k} \lambda}, \widehat{\Pi}_{\mathbf{k}^{\prime} \lambda^{\prime}}\right]=i \hbar \delta_{\lambda \lambda^{\prime}} \delta_{\mathbf{k} \mathbf{k}^{\prime}} \tag{7.70}
\end{equation*}
$$

exactly as for coordinates and momenta in earlier chapters. The dynamics is then given by the quantum version of the above energy which becomes the Hamiltonian

$$
\begin{equation*}
\widehat{H}=\sum_{\mathbf{k} \lambda}\left[\frac{1}{2 \epsilon_{0}} \widehat{\Pi}_{\mathbf{k} \lambda}^{\dagger} \widehat{\Pi}_{\mathbf{k} \lambda}+\frac{\epsilon_{0}}{2} \omega_{\mathbf{k}}^{2} \widehat{A}_{\mathbf{k} \lambda}^{\dagger} \widehat{A}_{\mathbf{k} \lambda}\right] \tag{7.71}
\end{equation*}
$$

similar to the harmonic oscillator in higher dimensions. In fact, the field has exactly the same dynamics as an infinite set of free oscillators. By this promotion of the classical field variables to quantum operators, we say that the electromagnetic field has been quantized. The classical field (7.60) has become a quantum field operator $\widehat{\mathbf{A}}(\mathbf{x}, t)$. The field at two different points will have non-trivial commutators which can be calculated from the above fundamental commutators (7.70) between the mode operators.
As for the simple harmonic oscillator, we can now introduce ladder operators in order to facilitate the construction of eigenstates and energy eigenvalues in the most direct way. In analogy with the (4.68) for the position operator and (4.69) the momentum operator in Chapter 4, we now write the field mode as

$$
\begin{equation*}
\widehat{A}_{\mathbf{k} \lambda}=\sqrt{\frac{\hbar}{2 \omega_{\mathbf{k}} \epsilon_{0}}}\left(\widehat{a}_{\mathbf{k} \lambda}+\widehat{a}_{-\mathbf{k} \lambda}^{\dagger}\right) \tag{7.72}
\end{equation*}
$$

and the conjugate momentum mode as

$$
\begin{equation*}
\widehat{\Pi}_{\mathbf{k} \lambda}=i \sqrt{\frac{\hbar \omega_{\mathbf{k}} \epsilon_{0}}{2}}\left(\widehat{a}_{\mathbf{k} \lambda}^{\dagger}-\widehat{a}_{-\mathbf{k} \lambda}\right) \tag{7.73}
\end{equation*}
$$

The $-\mathbf{k}$ in the last term of these operators is needed in order to satisfy the requirement $\hat{A}_{\mathbf{k} \lambda}^{\dagger}=\widehat{A}_{-\mathbf{k} \lambda}$ which comes from the corresponding reality condition of the classical field. It is now straightforward to show that the canonical field commutators (7.70) are satisfied when these new ladder operators obey

$$
\begin{equation*}
\left[\widehat{a}_{\mathbf{k} \lambda}, \widehat{a}_{\mathbf{k}^{\prime} \lambda^{\prime}}\right]=0, \quad\left[\widehat{a}_{\mathbf{k} \lambda}, \widehat{a}_{\mathbf{k}^{\prime} \lambda^{\prime}}^{\dagger}\right]=\delta_{\lambda \lambda^{\prime}} \delta_{\mathbf{k k}^{\prime}} \tag{7.74}
\end{equation*}
$$

Thus $\widehat{a}_{\mathbf{k} \lambda}$ is a lowering operator and $\widehat{a}_{\mathbf{k} \lambda}^{\dagger}$ is a raising operator. They lower and raise the number of exitations or quanta in the mode $(\mathbf{k} \lambda)$. This is given by the eigenvalue $n_{\mathbf{k} \lambda}=$ $0,1,2, \ldots$ of corresponding number operator $\widehat{n}_{\mathbf{k} \lambda}=\widehat{a}_{\mathbf{k} \lambda}^{\dagger} \widehat{a}_{\mathbf{k} \lambda}$. It has the commutators

$$
\begin{equation*}
\left[\widehat{n}_{\mathbf{k} \lambda}, \widehat{a}_{\mathbf{k}^{\prime} \lambda^{\prime}}\right]=-\delta_{\lambda \lambda^{\prime}} \delta_{\mathbf{k} \mathbf{k}^{\prime}} \widehat{a}_{\mathbf{k} \lambda}, \quad\left[\widehat{n}_{\mathbf{k} \lambda}, \widehat{a}_{\mathbf{k}^{\prime} \lambda^{\prime}}^{\dagger}\right]=+\delta_{\lambda \lambda^{\prime}} \delta_{\mathbf{k} \mathbf{k}^{\prime}} \widehat{a}_{\mathbf{k} \lambda} \tag{7.75}
\end{equation*}
$$

in complete analogy with what we had for a single oscillator. With help of these new ladder operator we can now write the Hamiltonian (7.71) as

$$
\begin{equation*}
\widehat{H}=\frac{1}{2} \sum_{\mathbf{k} \lambda} \hbar \omega_{\mathbf{k}}\left(\widehat{a}_{\mathbf{k} \lambda}^{\dagger} \widehat{a}_{\mathbf{k} \lambda}+\widehat{a}_{\mathbf{k} \lambda} \widehat{a}_{\mathbf{k} \lambda}^{\dagger}\right)=\sum_{\mathbf{k} \lambda} \hbar \omega_{\mathbf{k}}\left(\widehat{n}_{\mathbf{k} \lambda}+\frac{1}{2}\right) \tag{7.76}
\end{equation*}
$$

From (4.78) in Chapter 4 we know that the states

$$
\begin{equation*}
\left|n_{\mathbf{k} \lambda}\right\rangle=\sqrt{\frac{1}{n_{\mathbf{k} \lambda}!}}\left(\widehat{a}_{\mathbf{k} \lambda}^{\dagger}\right)^{n_{\mathbf{k} \lambda}}|0\rangle \tag{7.77}
\end{equation*}
$$

are eigenstates of the number operator, i.e $\widehat{N}_{\mathbf{k} \lambda}\left|n_{\mathbf{k} \lambda}\right\rangle=n_{\mathbf{k} \lambda}\left|n_{\mathbf{k} \lambda}\right\rangle$. Here $|0\rangle$ is the very special state without any excitations or quanta. This empty state is called the vacuum state in quantum field theory and satisfies the fundamental requirement

$$
\begin{equation*}
\widehat{a}_{\mathbf{k} \lambda}|0\rangle=0 \tag{7.78}
\end{equation*}
$$

by definition for all modes $(\mathbf{k}, \lambda)$. As expected, when the ladder operators actin on an excited state (7.77), they give

$$
\begin{equation*}
\widehat{a}_{\mathbf{k} \lambda}\left|n_{\mathbf{k} \lambda}\right\rangle=\sqrt{n_{\mathbf{k} \lambda}}\left|n_{\mathbf{k} \lambda}-1\right\rangle \tag{7.79}
\end{equation*}
$$

and

$$
\begin{equation*}
\widehat{a}_{\mathbf{k} \lambda}^{\dagger}\left|n_{\mathbf{k} \lambda}\right\rangle=\sqrt{n_{\mathbf{k} \lambda}+1}\left|n_{\mathbf{k} \lambda}+1\right\rangle \tag{7.80}
\end{equation*}
$$

like the ordinary oscillator operators act in (4.75) and (4.76). Thus $\widehat{a}_{\mathbf{k} \lambda}$ removes a quant from the state $\left|n_{\mathbf{k} \lambda}\right\rangle$ while $\widehat{a}_{\mathbf{k} \lambda}^{\dagger}$ creates one. In quantum field theories like here for the electromagnetic field, these ladder operators are called annihilation and creation operators.

The complete eigenstate of the Hamiltonian (7.76) can now be written as

$$
\begin{equation*}
\left|n_{\mathbf{k}_{1} \lambda_{1}}, n_{\mathbf{k}_{2} \lambda_{2}} \cdots\right\rangle=\Pi_{\mathbf{k} \lambda}\left|n_{\mathbf{k} \lambda}\right\rangle \tag{7.81}
\end{equation*}
$$

where the product goes over all available modes in the volume occupied by the field. It has quantized energies given by the eigenvalues

$$
\begin{equation*}
E=\sum_{\mathbf{k} \lambda}\left(n_{\mathbf{k} \lambda}+\frac{1}{2}\right) \hbar \omega_{\mathbf{k}} \tag{7.82}
\end{equation*}
$$

We see that the energy is uniquely given by the number $n_{\mathbf{k} \lambda}$ of quanta in each mode $(\mathbf{k}, \lambda)$. For this reason they are often called occupation numbers. An arbitrary eigenstate of the quantized electromagnetic field is completely specified by giving the value of all the occupation numbers.

In the empty or vacuum state all the ocupation numbers are zero. The energy of this state is therefore

$$
\begin{equation*}
E_{0}=\frac{1}{2} \sum_{\mathbf{k} \lambda} \hbar \omega_{\mathbf{k}} \tag{7.83}
\end{equation*}
$$

Since all the mode energies $\hbar \omega_{\mathbf{k}}$ are positive and there is an infinite number of modes, we get the surprising result that the quantized vacuum energy is infinite! It is due to the quantum fluctuations of the electromagnetic field in the vaccuum state which are exactly like the non-zero fluctuation $\langle 0| \widehat{x}^{2}|0\rangle$ of the oscillator in its ground state. And in the field there is an infinity of such oscillators. Remember that in classical physics the vacuum doesn't contain anything and therefore we can take its energy to be zero. This is a very big and deep problem! We will come back to it later, but just say here that in most instances we only measure energy differences and then this infinite constant will cancel out. But still...

The first excited state contains one quantum with wavenumber and polarization $(\mathbf{k}, \lambda)$. It is decribed by the state vector $|\mathbf{k} \lambda\rangle=\widehat{a}_{\mathbf{k} \lambda}^{\dagger}|0\rangle$ and is seen from (7.82) to have the energy $E=\hbar \omega_{\mathrm{k}}$ measured with respect to the zero-point energy $E_{0}$. More formally, this follows as the eigenvalue of the Hamiltonian (7.76) since

$$
\widehat{H}|\mathbf{k} \lambda\rangle=\sum_{\mathbf{k}^{\prime} \lambda^{\prime}} \hbar \omega_{\mathbf{k}^{\prime}} \hat{a}_{\mathbf{k}^{\prime} \lambda^{\prime}}^{\dagger} \widehat{a}_{\mathbf{k}^{\prime} \lambda^{\prime}}|\mathbf{k} \lambda\rangle=\sum_{\mathbf{k}^{\prime} \lambda^{\prime}} \hbar \omega_{\mathbf{k}^{\prime}} \widehat{a}_{\mathbf{k}^{\prime} \lambda^{\prime}}^{\dagger}\left[\widehat{a}_{\mathbf{k}^{\prime} \lambda^{\prime}}, \widehat{a}_{\mathbf{k} \lambda}^{\dagger}\right]|0\rangle
$$

using the definition (7.78) of the vacuum state. With the last commutator in (7.74) this now becomes

$$
\widehat{H}|\mathbf{k} \lambda\rangle=\hbar \omega_{\mathbf{k}} \widehat{a}_{\mathbf{k} \lambda}^{\dagger}|0\rangle=\hbar \omega_{\mathbf{k}}|\mathbf{k} \lambda\rangle
$$

This electromagnetric quantum of energy $E=\hbar \omega_{\mathbf{k}}$ is the photon. It is a particle since it also has the momentum $\mathbf{p}=\hbar \mathbf{k}$ as we will show in a minute. Since mass and energy for a relativistic particle is related by the Einstein equation $E^{2}=p^{2} c^{2}+m^{2} c^{4}$, it is in fact a massless particle, i.e. with mass $m=0$. Since all the knowledge we have about the particle is contained in the state evctor $|\mathbf{k} \lambda\rangle$, there is at this stage no obvious way to say where the particle is situated, or if it is meaningful at all to expect a well-defined position for it as for a classical particle. This would be equivalent to trying to find a wavefunction for the photon as for a non-relativistic particle. But this massless particle is really different and does not have an ordinary wavefunction.
The momentum density of the electromagnetic field is given by Poynting's vector $\epsilon_{0} \mathbf{E} \wedge \mathbf{B}$. Using again the mode expansion and integrating over space, one thus finds the total momentum in the field as

$$
\mathbf{P}=-i \epsilon_{0} \sum_{\mathbf{k}} \dot{\mathbf{A}}_{\mathbf{k}}^{*} \wedge\left(\mathbf{k} \wedge \mathbf{A}_{\mathbf{k}}\right)
$$

Here we recognize $\boldsymbol{\Pi}_{\mathbf{k}}=\epsilon_{0} \dot{\mathbf{A}}_{\mathbf{k}}^{*}$ as the canonical momentum of the field. Using now the triple vector product formula $\mathbf{A} \wedge(\mathbf{B} \wedge \mathbf{C})=(\mathbf{A} \cdot \mathbf{C}) \mathbf{B}-(\mathbf{A} \cdot \mathbf{B}) \mathbf{C}$ and the transversality condition $\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}}=0$ from the Coulomb gauge condition, one obtains

$$
\mathbf{P}=-i \sum_{\mathbf{k}}\left(\boldsymbol{\Pi}_{\mathbf{k}} \cdot \mathbf{A}_{\mathbf{k}}\right) \mathbf{k}
$$

As expected, the momentum of the mode with wavenumber $\mathbf{k}$ is along that same direction. Decomposing these vectors onto the two polarization vectors and quantizing by replaceming of the classical field variables with the quantum operator (7.72) and (7.73), we find

$$
\widehat{\mathbf{P}}=\frac{1}{2} \sum_{\mathbf{k} \lambda} \hbar \mathbf{k}\left(\widehat{a}_{\mathbf{k} \lambda}^{\dagger} \widehat{a}_{\mathbf{k} \lambda}+\widehat{a}_{\mathbf{k} \lambda}^{\dagger} \widehat{a}_{-\mathbf{k} \lambda}^{\dagger}-\widehat{a}_{-\mathbf{k} \lambda} \widehat{a}_{\mathbf{k} \lambda}-\widehat{a}_{-\mathbf{k} \lambda} \widehat{a}_{-\mathbf{k} \lambda}^{\dagger}\right)
$$

The two middle terms are odd under $\mathbf{k} \rightarrow-\mathbf{k}$ and thus each gives zero in the sum. With the same shift of summation variable in the last term combined with a commutation, it gives the same as the first. We therefore have the simple result

$$
\begin{equation*}
\widehat{\mathbf{P}}=\sum_{\mathbf{k} \lambda} \hbar \mathbf{k} \widehat{a}_{\mathbf{k} \lambda}^{\dagger} \widehat{a}_{\mathbf{k} \lambda} \tag{7.84}
\end{equation*}
$$

for the momentum operator. This explains why the one-photon eigenstate $|\mathbf{k} \lambda\rangle$ not only has the energy eigenvalue $E=\hbar \omega_{\mathbf{k}}$ but also the momentum eigenvalue $\mathbf{p}=\hbar \mathbf{k}$. Similarly, the two-photon state $\left|\mathbf{k} \lambda, \mathbf{k}^{\prime} \lambda^{\prime}\right\rangle$ has the energy $E=\hbar \omega_{\mathbf{k}}+\hbar \omega_{\mathbf{k}^{\prime}}$ and the momentum $\mathbf{p}=\hbar \mathbf{k}+\hbar \mathbf{k}^{\prime}$.

One can also derive an operator which gives the spin of a photon state. Instead of going through the derivation, it is simpler to note that the electromagnetic field is decribed by the vector $\mathbf{A}(x)$. It will then as all other 3 -vectors, transform as a spin- 1 object under rotations. The photon thus has spin $S=1$. But since it is massless and thus does not have any restframe, the spin has only two components. These correspond to the spin pointing along or opposite to the direction of the wavevector $\mathbf{k}$. The component of the spin of a massless particle in this direction is called the helicity of the particle and thus can take the values $\lambda= \pm 1$ for the photon. It is intimately connected with the polarization state of a classical, monochromatic electromagnetic wave in the same direction and with the same frequency $\omega=k c$. If the motion is along the $z$-axis, the two polarization vectors

$$
\begin{equation*}
\mathbf{e}_{R / L}=\sqrt{\frac{1}{2}}\left(\mathbf{e}_{x} \pm i \mathbf{e}_{y}\right) \tag{7.85}
\end{equation*}
$$

correspond to the electric field vector $\mathbf{E}$ (or magnetic field $\mathbf{B}$ ) rotationg around the $z$ axis in the positive, respectively negative sense with frequency $\omega$. Thus helicity $\lambda=+1$ corresponds to right-handed polarization and helicity $\lambda=-1$ corresponds to left-handed polarization
What remains is to fully write out the electromagnetic field operator which follows from the expansion (7.60) when the modes $\mathbf{A}_{\mathbf{k}}(t)$ are replaced by the quantum operators (7.72). Writing these in the Heisenberg picture as for each oscillator mode like (4.86) in Chapter 4, one gets

$$
\begin{align*}
\widehat{\mathbf{A}}(\mathbf{x}, t) & =\sum_{\mathbf{k} \lambda} \sqrt{\frac{\hbar}{2 \omega_{\mathbf{k}} \epsilon_{0} V}} \mathbf{e}_{\lambda}\left[\widehat{a}_{\mathbf{k} \lambda} e^{-i \omega_{\mathbf{k}} t}+\widehat{a}_{-\mathbf{k} \lambda}^{\dagger} e^{i \omega_{\mathbf{k}} t}\right] e^{i \mathbf{k} \cdot \mathbf{x}} \\
& =\sum_{\mathbf{k} \lambda} \sqrt{\frac{\hbar}{2 \omega_{\mathbf{k}} \epsilon_{0} V}}\left[\mathbf{e}_{\lambda} \widehat{a}_{\mathbf{k} \lambda} e^{i\left(\mathbf{k} \cdot \mathbf{x}-\omega_{\mathbf{k}} t\right)}+\mathbf{e}_{\lambda}^{*} \widehat{a}_{\mathbf{k} \lambda}^{\dagger} e^{-i\left(\mathbf{k} \cdot \mathbf{x}-\omega_{\mathbf{k}} t\right)}\right] \tag{7.86}
\end{align*}
$$

after making the replacement $\mathbf{k} \rightarrow-\mathbf{k}$ in the last sum. Notice that we then have $\mathbf{e}_{\lambda}(-\mathbf{k}) \rightarrow \mathbf{e}_{\lambda}^{*}(\mathbf{k})$ since the polarization vectors can in general be complex, corresponding to circular polarization. The first term can be interpreted as annihilating a photon at position $\mathbf{x}$ at time $t$ while the last term describes the creation of a photon at the same spacetime point. These localized photons are seen to be infinite superpositions of photons with well-defined wavenumbers or momenta, as expected from the Heisenberg uncertainty relations.

Each term of the electromagnetic qyantum field operator (7.86) has the same structure as the classical wave (7.43) or mode with a given wavevector. An obvious question is now what the correspondence is between the above field operator describing annihilation and creation of a single photon, and a classical wave. From the correspondence principle we know that a quantum system can be described in classical terms when its quantum numbers become large. In our case that corresponds to having a large number of photons in the system. A classical, electromagnetic wave can then be thought of as made up of many photons in a particular way. In fact, from the existence of coherent states for the harmonic oscillator in Chapter 4, we know that in a coherent state of photons with a definite wavenumber and polarization, the expectation value of the quantum field operator will be exactly a classical wave. But what is then the static Coulomb potential in terms of photons? Apparently there must be more to quantum electromagnetism than just these massless field quanta. The full theory of quantum electrodynamics (QED) answers these remaining questions in a more satisfactory way.

### 7.9 Spontaneous emission

Let us now return to the process of emission of light from an excited state of an atom illustrated in Fig. 7.5. When the electromagnetic field now is quantized, we can calculate the rate for emission of one or several photons. The simplest process is the emission of just one photon which we will consider here. Nothing has changed in the interaction Hamiltonian (7.42) except for the vector potential now being a quantum field operator. In lowest order perturbation theory the transition matrix element is therefore

$$
\begin{equation*}
V_{f i}=\frac{e}{m_{e}}\langle f| \widehat{\mathbf{p}} \cdot \widehat{\mathbf{A}}(\widehat{\mathbf{x}}, 0)|i\rangle \tag{7.87}
\end{equation*}
$$

In the argument of the field operator we find the position operator $\widehat{\mathbf{x}}$ of the electron since it is there the field acts.
We will consider an initial state of $n_{\mathbf{k} \lambda}$ photons in the mode $(\mathbf{k} \lambda)$ and the atom in the initial state $|m\rangle$. The full initial state can therefore be written as $|i\rangle=\left|m ; n_{\mathbf{k} \lambda}\right\rangle$. For the emission of one photon, it is the first part of the field operator (7.86) which will contribute. It creates one additional photon in the same mode so that the final state of the coupled system is $|f\rangle=\left|n ; n_{\mathbf{k} \lambda}+1\right\rangle$ when the atom itself is found in the state $|n\rangle$. Inserting for the field operator, we thus have for the transition matrix element

$$
V_{f i}=\frac{e}{m_{e}} \sqrt{\frac{\hbar}{2 \omega_{\mathbf{k}} \epsilon_{0} V}}\langle n| e^{-i \mathbf{k} \cdot \widehat{\mathbf{x}}} \mathbf{e}_{\lambda}^{*} \cdot \hat{\mathbf{p}}|m\rangle\left\langle n_{\mathbf{k} \lambda}+1\right| \widehat{a}_{\mathbf{k} \lambda}^{\dagger}\left|n_{\mathbf{k} \lambda}\right\rangle
$$

The photonic matrix element is simply $\sqrt{n_{\mathbf{k} \lambda}+1}$ from (7.80) and the atomic part is exactly the same as in the previous, semi-classical case (7.51). In the dipole approximation we can use (7.46) to write the matrix element $\mathbf{p}_{n m}=-i m_{e} \omega_{\mathbf{k}} \mathbf{x}_{n m}$. With help of the Fermi Golden Rule, we thus get for the transition rate

$$
\begin{equation*}
\Gamma_{n m}^{e m}=\frac{2 \pi}{\hbar}\left(\frac{e}{m_{e}}\right)^{2} \frac{\hbar}{2 \omega_{\mathbf{k}} \epsilon_{0} V}\left(m_{e} \omega_{\mathbf{k}}\right)^{2}\left(n_{\mathbf{k} \lambda}+1\right)\left|\mathbf{e}_{\lambda}^{*} \cdot \mathbf{x}_{n m}\right|^{2} \delta\left(E_{n}-E_{m}+\hbar \omega_{\mathbf{k}}\right) \tag{7.88}
\end{equation*}
$$

The total rate is now obtained by summing over all wavenumbers $\mathbf{k}$ and the two polarization vector $\mathbf{e}_{\lambda}$ of the final photon. In the large-volume limit we write the first sum as in (7.62), giving for the total transition rate

$$
\begin{equation*}
\Gamma_{n m}^{e m}=\frac{e^{2} \pi}{\epsilon_{0}} \sum_{\lambda} \int \frac{d^{3} k}{(2 \pi)^{3}}\left|\mathbf{e}_{\lambda}^{*} \cdot \mathbf{x}_{n m}\right|^{2} \omega_{\mathbf{k}}\left(n_{\mathbf{k} \lambda}+1\right) \delta\left(E_{n}-E_{m}+\hbar \omega_{\mathbf{k}}\right) \tag{7.89}
\end{equation*}
$$

where $d^{3} k=k^{2} d k d \Omega_{\mathbf{k}}$. The integration over the photon frequency $\omega_{\mathbf{k}}=k c=\omega$ fixes it to the value $\omega=\left(E_{m}-E_{n}\right) / \hbar$. Therefore

$$
\Gamma_{n m}^{e m}=\frac{e^{2} \omega^{3}}{8 \pi^{2} \epsilon_{0} \hbar c^{3}} \sum_{\lambda} \int d \Omega_{\mathbf{k}}\left|\mathbf{e}_{\lambda}^{*} \cdot \mathbf{x}_{n m}\right|^{2}\left(n_{\mathbf{k} \lambda}+1\right)
$$

For the combined integration over photon directions and polarization sum, we proceed as in the derivation of (7.48), i.e.

$$
\sum_{\lambda} \int d \Omega_{\mathbf{k}}\left|\mathbf{e}_{\lambda}^{*} \cdot \mathbf{x}_{n m}\right|^{2}=8 \pi \cdot \frac{1}{3}\left|\mathbf{x}_{n m}\right|^{2}
$$

since each of the two polarization directions give the same. But we can also now obtain this result by first summing over the two polarization vectors using the relation (7.67),

$$
\begin{equation*}
\sum_{\lambda}\left|\mathbf{e}_{\lambda}^{*} \cdot \mathbf{x}_{n m}\right|^{2}=\left|\mathbf{x}_{n m}\right|^{2}-\left|\widehat{\mathbf{k}} \cdot \mathbf{x}_{n m}\right|^{2} \tag{7.90}
\end{equation*}
$$

where again $\widehat{\mathbf{k}}$ is a unit vector along $\mathbf{k}$. If this makes the angle $\theta$ with $\mathbf{x}_{n m}$ considered as a fixed vector, we thus get by integration over all directions

$$
\begin{equation*}
\sum_{\lambda} \int d \Omega_{\mathbf{k}}\left|\mathbf{e}_{\lambda}^{*} \cdot \mathbf{x}_{n m}\right|^{2}=2 \pi\left|\mathbf{x}_{n m}\right|^{2} \int_{0}^{\pi} d \theta \sin \theta\left(1-\cos ^{2} \theta\right)=\frac{8 \pi}{3}\left|\mathbf{x}_{n m}\right|^{2} \tag{7.91}
\end{equation*}
$$

which is seen to agree with what we found above.
Including now all the factors and rearranging, we thus have for the full transition rate

$$
\begin{equation*}
\Gamma_{n m}^{e m}=\frac{e^{2} \omega^{3}}{8 \pi^{2} \epsilon_{0} \hbar c^{3}} \frac{8 \pi}{3}\left|\mathbf{x}_{n m}\right|^{2}\left(n_{\mathbf{k}}+1\right) \tag{7.92}
\end{equation*}
$$

where $n_{\mathbf{k}}$ is the number of photons in each polarization mode. For thermal radiation it is given by the Bose-Einstein distribution function

$$
\begin{equation*}
n_{\mathbf{k}}=\frac{1}{e^{\beta \hbar \omega_{\mathbf{k}}}-1} \tag{7.93}
\end{equation*}
$$

We see that the total emission rate splits into two parts. The first part is proportional to the number of photons in the initial state. One says that the emission is induced by these photons. The corresponding induced emission rate can then be written as

$$
\begin{equation*}
\Gamma_{n m}^{i n d u c}=\frac{4 \pi^{2} \alpha c}{3 \hbar}\left|\mathbf{x}_{n m}\right|^{2} u(\omega) \tag{7.94}
\end{equation*}
$$

where $u(\omega)=\left(\omega^{2} / \pi^{2} c^{3}\right) \hbar \omega n_{\mathbf{k}}$ as for the Planck distribution (7.55). This is seen to be in complete agreement with the semi-classical result (7.52), derived in the limit where the electromagnetic field could be treated classically.

But there is an additional term arising from the correct quantum treatment of the field given here. From the full rate (7.92), setting the number $n_{\mathbf{k}}$ of initial photons equal to zero, we get the spontaneous emission rate. It is seen to be

$$
\begin{equation*}
\Gamma_{n m}^{s p o n t}=\frac{4}{3} \alpha \omega\left|\frac{\omega}{c} \mathbf{x}_{n m}\right|^{2} \tag{7.95}
\end{equation*}
$$

and is also in agreement with what followed from the Einstein equilibrium argument in (7.58).

We will now calculate this spontaneous emission rate for the simplest transition in the hydrogen atom from the first excited state $|2 p\rangle$ to the ground state $|1 s\rangle$. The matrix element $\mathbf{x}_{n m}$ is most simply calculated by expressing the electron position operator $\widehat{\mathbf{x}}$ in terms of the spherical components $\widehat{x}_{ \pm}=(\widehat{x} \pm i \widehat{y}) / \sqrt{2}$ so that

$$
\begin{align*}
\left|\mathbf{x}_{n m}\right|^{2} & =|\langle 1 s| \widehat{\mathbf{x}}| 2 p\rangle\left.\right|^{2} \\
& \left.\left.\left.=\left|\langle 1 s| \widehat{x}_{-}\right| 2 p\right\rangle\left.\right|^{2}+\left|\langle 1 s| \widehat{x}_{+}\right| 2 p\right\rangle\left.\right|^{2}+|\langle 1 s| \widehat{z}| 2 p\right\rangle\left.\right|^{2} \tag{7.96}
\end{align*}
$$

In the coordinate representation the corresponding wavefunctions $\psi_{n l m}(r, \theta, \phi)$ are

$$
\psi_{100}=R_{1 s}(r) Y_{00}(\theta, \phi), \quad \psi_{21 m}=R_{2 p}(r) Y_{1 m}(\theta, \phi)
$$

where the radial functions were calculated in Chapter 5. We can then also write

$$
\begin{equation*}
x_{ \pm}=\sqrt{\frac{1}{2}} r \sin \theta e^{ \pm i \phi}=\mp \sqrt{\frac{4 \pi}{3}} r Y_{1, \pm 1}(\theta, \phi) \tag{7.97}
\end{equation*}
$$

and $z=r \cos \theta=\sqrt{4 \pi / 3} r Y_{10}(\theta, \phi)$. Since now $Y_{00}=1 / \sqrt{4 \pi}$ is just a constant, we see from the orthogonality of the spherical harmonics in (5.96) that the matrix element $\langle 1 s| \widehat{x}_{-}|2 p\rangle$ will be non-zero only when the atom is in the initial $2 p$-state with $L_{z}$ eigenvalue equal to $m=+1$. It will then take the value

$$
\begin{align*}
\langle 1 s| \widehat{x}_{-}|2 p,+1\rangle & =-\sqrt{\frac{1}{4 \pi}} \int d^{3} r R_{1 s}^{*} r \sqrt{\frac{4 \pi}{3}} Y_{1,+1}^{*} Y_{1,+1} R_{2 p} \\
& =-\sqrt{\frac{1}{3}} \int_{0}^{\infty} d r r^{3} R_{1 s} R_{2 p} \int d \Omega Y_{1,+1}^{*} Y_{1,+1}=-\sqrt{\frac{1}{3}} I \tag{7.98}
\end{align*}
$$

since the angular integral is one. The radial integration gives

$$
\begin{align*}
I=\int_{0}^{\infty} d r r^{3} R_{1 s} R_{2 p} & =\sqrt{\frac{4}{a_{0}^{3}}} \sqrt{\frac{1}{24 a_{0}^{3}}} \int_{0}^{\infty} d r r^{3} e^{-r / 2 a_{0}} \frac{r}{a_{0}} e^{-r / a_{0}} \\
& =a_{0} \sqrt{\frac{1}{6}}\left(\frac{4}{3}\right)^{4} \tag{7.99}
\end{align*}
$$

If we instead had considered the atom to be intially in the state with $m=-1$, then the middle term in (7.96) would have contributed with the same value. Similarly, the last term would have contributed also the same if the atom was initially in the state with $m=0$. When these three initial states have the same probability, we should take the average, i.e. the sum of the three terms in (7.96) and divide by three. This is then the same as (7.99) again.

Besides the matrix element (7.98), the transition rate (7.95) also involves the frequency $\omega$ of the emitted photon. Here it has the value

$$
\begin{equation*}
\hbar \omega=E_{2 p}-E_{1 s}=-\frac{1}{2} \alpha^{2} m_{e} c^{2}\left(\frac{1}{2^{2}}-\frac{1}{1^{2}}\right)=\frac{3}{8} \alpha^{2} m_{e} c^{2} \tag{7.100}
\end{equation*}
$$

which gives for the rate

$$
\begin{equation*}
\Gamma(2 p \rightarrow 1 s+\gamma)=\frac{4 \alpha \omega^{3}}{3 c^{2}} \frac{1}{3} I^{2}=\alpha^{5}\left(\frac{2}{3}\right)^{8} \frac{m_{e} c^{2}}{\hbar} \tag{7.101}
\end{equation*}
$$

Putting in numbers, we find $\Gamma=4.13 \times 10^{-7} \mathrm{eV} / \hbar$ corresponding to the lifetime $\tau=$ $1 / \Gamma=1.61 \times 10^{-9}$ sec. It is short as already mentioned, but much longer than many unstable nuclear states and elementary particles.

What about the other excited state $|2 s\rangle$ ? It would have a one-photon decay to the ground state given by the matrix element $\langle 1 s| \widehat{\mathbf{x}}|2 s\rangle$. But writing it out, we see that it is zero from the angular integration. More generally, since both $|1 s\rangle$ and $|2 s\rangle$ have wavefunctions that are rotationally invariant and with even parity, and $\widehat{\mathbf{x}}$ is a vector operator and therefore has odd parity, the matrix element must be zero. The same argument holds even if we had not done the dipole appoximation. Thus the transition rate is exactly zero in first order perturbation theory. But in second order we must also include the two-photon operator stemming from the second term in the perturbation (7.86) and we will get a non-zero result. Since it involves more powers of the finestructure constant, the transition rate will be smaller and the lifetime of this state correspondingly longer. In fact, a rather difficult calculation gives a lifetime of around eight seconds. It is almost an eternity compared with the lifetime for the $|2 p\rangle$ state.

### 7.10 Angular distributions

Having now established a quantum theory of photons, we can ask more detailed questions about the what happens in the decay. For instance, if the atom was initially in the $|2 p\rangle$ with its angular momentum along the $z$-axis so that the azimuthal quantum number $m=+1$, what is then the probability for observing the photon from the decay coming out in a certain direction? Or we can be even more specific, what is the probability to observe a right-handed photon in this direction? We can then no longer use the averaged transition rate (7.95), but must use the more basic formula (7.88) where we only consider the spontaneous part.

If the created photon has polarization (or helcity) $\lambda$ with wavenumber $\mathbf{k}$ in the direction $\Omega_{\mathbf{k}}$ and forming the angle $\beta$ with the $z$-axis as shown in the Fig.7.7, then we have for the differential transition rate

$$
d \Gamma=\pi e^{2} \frac{d^{3} k}{(2 \pi)^{3}} \frac{\omega}{\epsilon_{0}}\left|\mathbf{e}_{\lambda}^{*} \cdot \mathbf{x}_{n m}\right|^{2} \delta\left(E_{n}-E_{m}+\hbar \omega\right)
$$

We have here assumed that the quantization volume is so large that we can replace the sum over modes with an integral according to the general prescription (7.62) with $d^{3} k=k^{2} d k d \Omega_{\mathbf{k}}$. Integrating over all frequencies $\omega=k c$, we then again fix the frequency


Figure 7.7: Photon with wavevector $\mathbf{k}$ being emitted in the $x z$-plane. Initial atomic spin is along the $z$-axis.
to be $\omega=\left(E_{m}-E_{n}\right) / \hbar$ and get

$$
\begin{equation*}
\frac{d \Gamma}{d \Omega}=\frac{\alpha \omega^{3}}{2 \pi c^{2}}\left|\mathbf{e}_{\lambda}^{*} \cdot \mathbf{x}_{n m}\right|^{2} \tag{7.102}
\end{equation*}
$$

Choosing the wavenumber vector to be in the $x z$-plane as shown in the figure so that $\mathbf{k}=k\left(\mathbf{e}_{z} \cos \beta+\mathbf{e}_{x} \sin \beta\right)$, the two linear and independent polarization vectors can be taken to be $\mathbf{e}_{1}=-\mathbf{e}_{z} \sin \beta+\mathbf{e}_{x} \cos \beta$ and $\mathbf{e}_{2}=\mathbf{e}_{y}$. Notice that defining conditions $\mathbf{k} \cdot \mathbf{e}_{1}=\mathbf{k} \cdot \mathbf{e}_{2}=0$ both are satisfied. A right-handed photon in this direction thus has the polarization vector $\mathbf{e}_{R}=\left(\mathbf{e}_{1}+i \mathbf{e}_{2}\right) / \sqrt{2}$ so that

$$
\begin{equation*}
\mathbf{e}_{R}^{*} \cdot \widehat{\mathbf{x}}=\sqrt{\frac{1}{2}}(-\widehat{z} \sin \beta+\widehat{x} \cos \beta-i \widehat{y}) \tag{7.103}
\end{equation*}
$$

We then use (7.97) to express the components of the vector $\mathbf{x}$ in terms of the three spherical harmonics $Y_{1 \pm}$ and $Y_{10}$. The transition amplitude for the emission of a righthanded photon from the atom in the initial state $|2 p,+1\rangle$ is then

$$
\begin{aligned}
& \langle 1 s| \mathbf{e}_{R}^{*} \cdot \widehat{\mathbf{x}}|2 p,+1\rangle=\sqrt{\frac{1}{6}} \int_{0}^{\infty} d r r^{3} R_{1 s} R_{2 p} \times \\
& \int d \Omega\left[-Y_{10} \sin \beta+\sqrt{\frac{1}{2}}\left(-Y_{1,+1}+Y_{1,-1}\right) \cos \beta+\sqrt{\frac{1}{2}}\left(Y_{1,+1}+Y_{1,-1}\right)\right] Y_{1,+1}
\end{aligned}
$$

in analogy to (7.98). The angular integration picks out a non-zero contribution from $\int d \Omega Y_{1,-1} Y_{1,+1}=-1$ since $Y_{1,-1}=-Y_{1,+1}^{*}$. It gives the result

$$
\begin{equation*}
\langle 1 s| \mathbf{e}_{R}^{*} \cdot \widehat{\mathbf{x}}|2 p,+1\rangle=-\sqrt{\frac{1}{12}} I(1+\cos \beta) \tag{7.104}
\end{equation*}
$$

where the factor $I$ is again the angular integral (7.99). Substituting this result into (7.102) with the frequency (7.100), we find the angular distribution

$$
\begin{equation*}
\frac{d \Gamma}{d \Omega}\left(2 p_{+1} \rightarrow 1 s+\gamma_{R}\right)=\frac{m_{e} c^{2} \alpha^{5}}{2 \pi \hbar}\left(\frac{2}{3}\right)^{6} \frac{1}{12}(1+\cos \beta)^{2} \tag{7.105}
\end{equation*}
$$

for the emission of a positive-helicity photon. It is zero when $\beta=\pi$ when the photon moves in the negative $z$-direction. This is due to spin conservation since the initial, atomic spin of $m=+1$ was along the positive $z$-direction.

An almost identical calculation also gives the result

$$
\begin{equation*}
\frac{d \Gamma}{d \Omega}\left(2 p_{+1} \rightarrow 1 s+\gamma_{L}\right)=\frac{m_{e} c^{2} \alpha^{5}}{2 \pi \hbar}\left(\frac{2}{3}\right)^{6} \frac{1}{12}(1-\cos \beta)^{2} \tag{7.106}
\end{equation*}
$$

when the emitted photon is coming out in the same direction, but now with negative helecity. This transition rate is now zero when the photon comes out along the positive $z$ direction, again from spin conservation. We if now don't observe the photon polarization, we must add these two partial results to get the full differential transition rate. One then obtains

$$
\begin{equation*}
\frac{d \Gamma}{d \Omega}\left(2 p_{+1} \rightarrow 1 s+\gamma\right)=\frac{m_{e} c^{2} \alpha^{5}}{2 \pi \hbar}\left(\frac{2}{3}\right)^{6} \frac{1}{6}\left(1+\cos ^{2} \beta\right) \tag{7.107}
\end{equation*}
$$

If we now don't even observe the direction of the out-going photon, we must integrate over all photon directions to obtain the full emision rate. It becomes

$$
\begin{align*}
\Gamma\left(2 p_{+1} \rightarrow 1 s+\gamma\right) & =\frac{m_{e} c^{2} \alpha^{5}}{\hbar}\left(\frac{2}{3}\right)^{6} \frac{1}{6} \int_{0}^{\pi} d \beta \sin \beta\left(1+\cos ^{2} \beta\right) \\
& =\frac{m_{e} c^{2} \alpha^{5}}{\hbar}\left(\frac{2}{3}\right)^{8} \tag{7.108}
\end{align*}
$$

Needless to say, it is in full agreement with the result (7.101) from the previous calculation.

The transition amplitudes for emission of photons from the other states like $|2 p, 0\rangle$ and $|2 p,-1\rangle$ can be read out directly from (7.104). For instance, the amplitude for emitting a right-handed photon from $|2 p, 0\rangle$ is seen to be $\alpha \sin \beta$ while the amplitude from $|2 p,-1\rangle$ is the same as for a left-handed photon from $|2 p,+1\rangle$ leading to the same angular distribution as in (7.106). In fact, we can write down the results immediately by using the rotation matrices in Chapter 5 . Using the fact that the photon is a spin- 1 particle with helicity $\lambda= \pm 1$ which is the value of the spin along its wavevector in direction $\beta$ from the $z$-axis and the initial atom is in the state $|2 p, m\rangle$, the amplitude for the emission $|2 p, m\rangle \rightarrow|2 s\rangle+\gamma_{\lambda}$ has an angular dependency simply given by the rotation matrix element $\propto d_{m \lambda}^{(1)}(\beta)$. This is easily verified by comparing the above results with the entries in the spin-1 rotation matrix (5.165).

### 7.11 Magnetic transitions

In the above we have considered electric transitions which doesn't involve the intrinsic spin of the electron. But from the Pauli-equation (7.30) there results also a coupling
$\widehat{H}_{B}=-\widehat{\boldsymbol{\mu}} \cdot \mathbf{B}$ where the magnetic moment operator is $\boldsymbol{\mu}=(e \hbar / 2 m) \boldsymbol{\sigma}$ when expressed in terms of the Pauli matrices. This will induce magnetic transitions between atomic levels.

Let us now consider such a transition $|i\rangle \rightarrow|f\rangle+\gamma$ in the hydrogen atom. The lowest order transition matrix element is then

$$
\begin{equation*}
V_{f i}=-\frac{e \hbar}{2 m_{e}}\langle f| \boldsymbol{\sigma} \cdot \mathbf{B}(\widehat{\mathbf{x}})|i\rangle \tag{7.109}
\end{equation*}
$$

Since the magnetic field $\mathbf{B}=\boldsymbol{\nabla} \wedge \mathbf{A}$ and denoting the photon wave vector by $\mathbf{k}$, we find

$$
V_{f i}=\frac{i e \hbar}{2 m_{e}} \sqrt{\frac{\hbar}{2 \omega \epsilon_{0} V}}\langle f| \boldsymbol{\sigma} \cdot\left(\mathbf{k} \wedge \mathbf{e}_{\lambda}^{*}\right) e^{-i \mathbf{k} \cdot \widehat{\mathbf{x}}}|i\rangle
$$

as we did in the previous chapter for spontaneous emission caused by the usual elctric interaction. For optical transitions we are again allowed to make the dipole approximation so we can drop the exponential term in the matrix element. From the Golden Rule we then have the transition rate

$$
\Gamma=\frac{2 \pi}{\hbar}\left(\frac{e \hbar}{2 m_{e}}\right)^{2} \frac{\hbar}{2 \omega \epsilon_{0} V}\left|\mathbf{e}_{\lambda}^{*} \cdot\left(\mathbf{k} \wedge \boldsymbol{\sigma}_{f i}\right)\right|^{2} \delta\left(E_{f}-E_{i}+\hbar \omega\right)
$$

where now $\boldsymbol{\sigma}_{f i}=\langle f| \boldsymbol{\sigma}|i\rangle$ is an atomic spin matrix element. The total transition rate with the photon within $d^{3} k=k^{2} d k d \Omega$ thus becomes

$$
\begin{align*}
\Gamma & =\frac{2 \pi}{\hbar}\left(\frac{e \hbar}{2 m_{e}}\right)^{2} \sum_{\lambda} \int \frac{d^{3} k}{(2 \pi)^{3}} \frac{\hbar}{2 \omega \epsilon_{0}}\left|\mathbf{e}_{\lambda}^{*} \cdot\left(\mathbf{k} \wedge \boldsymbol{\sigma}_{f i}\right)\right|^{2} \delta\left(E_{f}-E_{i}+\hbar \omega\right) \\
& =\left(\frac{e \hbar}{2 m_{e}}\right)^{2} \frac{\omega}{8 \pi^{2} \epsilon_{0} \hbar c^{3}} \sum_{\lambda} \int d \Omega\left|\mathbf{e}_{\lambda}^{*} \cdot\left(\mathbf{k} \wedge \boldsymbol{\sigma}_{f i}\right)\right|^{2} \tag{7.110}
\end{align*}
$$

Summing over the two polarization vectors $\mathbf{e}_{\lambda}$ using (7.67) as in (7.90), gives

$$
\sum_{\lambda}\left|\mathbf{e}_{\lambda}^{*} \cdot\left(\mathbf{k} \wedge \boldsymbol{\sigma}_{f i}\right)\right|^{2}=\left|\mathbf{k} \wedge \boldsymbol{\sigma}_{f i}\right|^{2}-\left|\widehat{\mathbf{k}} \cdot\left(\mathbf{k} \wedge \boldsymbol{\sigma}_{f i}\right)\right|^{2}
$$

where the the last term is zero since $\widehat{\mathbf{k}} \wedge \mathbf{k}=0$. After having integrated over all photon directions, we are then left with

$$
\sum_{\lambda} \int d \Omega\left|\mathbf{k} \wedge \boldsymbol{\sigma}_{f i}\right|^{2}=k^{2} \int d \Omega\left(\left|\boldsymbol{\sigma}_{f i}\right|^{2}-\left(\widehat{\mathbf{k}} \cdot \boldsymbol{\sigma}_{f i}\right)^{2}\right)=\frac{8 \pi}{3} k^{2}\left|\boldsymbol{\sigma}_{f i}\right|^{2}
$$

as in the previous derivation of (7.91). Using this in the above expresion for the rate, we thus get the final result

$$
\begin{equation*}
\Gamma=\left(\frac{e \hbar}{2 m_{e}}\right)^{2} \frac{\omega}{8 \pi^{2} \epsilon_{0} \hbar c^{3}} \frac{8 \pi}{3} k^{2}\left|\boldsymbol{\sigma}_{f i}\right|^{2}=\frac{4}{3} \alpha \omega\left|\frac{\hbar \omega}{2 m_{e} c^{2}} \boldsymbol{\sigma}_{f i}\right|^{2} \tag{7.111}
\end{equation*}
$$

It should be compared with the spontaneous rate (7.95) for electric transitions. They take place between states with different values of the orbital angular momentym. But since this magnetic rate involves matrix element of the intrinsic spin $\boldsymbol{\sigma}$, it describes transitions btween states where this spin changes. We will come back to it in the next chapter when we consider the hyperfine splitting in atoms.

### 7.12 Photo-electric effect

We have just seen how to calculate the rate for photon absorption where the electron bound in an atom is raised to an excited level. If the photon energy $\hbar \omega$ is sufficiently high, the excited electron gets so much energy that it can escape the ionized atom as a free particle. This is the photo-electric effect. The simplest version is for a hydrogen-like atom initially in the ground state with energy $E_{0}=-\alpha^{2} Z^{2} m_{e} c^{2} / 2$ when the nucleus has the charge $Z e$. For the ejected electron in the final state to be free from the Coulomb attraction to the remaining ion, its kinetic energy $E_{f}=p^{2} / 2 m_{e}=\hbar \omega+E_{0}$ must be much larger than the ionization energy $-E_{0}$. This implies that the velocity $v=p / m_{e}$ of the final electron is sufficiently high so that the constraint $Z \alpha / \beta \ll 1$ with $\beta=v / c$, is satisfied. Then the final electron momentum follows directly from the photon energy by $p^{2} / 2 m_{e}=\hbar \omega$. At the same time we also want this electron to be non-relativistic. For our calculation to be valid, the allowed photon energies must thus be within the range $13.6 \mathrm{eV} \ll \hbar \omega \ll 0.511 \mathrm{MeV}$.

The transition matrix element from a general, initial initial state $|i\rangle$ follows from the first term in the interaction Hamiltonian(7.42) as

$$
\begin{equation*}
V_{f i}=-\frac{e}{m_{e}} \sqrt{\frac{\hbar}{2 \omega \epsilon_{0} V}}\langle f| \mathbf{e}_{\lambda} \cdot \widehat{\mathbf{p}} e^{i \mathbf{k} \cdot \widehat{\mathbf{r}}}|i\rangle \tag{7.112}
\end{equation*}
$$

where $\mathbf{e}_{\lambda}$ is the polarization vector for the photon with wave vector $\mathbf{k}$. The electron in the final state has the wave function $\psi_{f}=e^{i \mathbf{p} \cdot \mathbf{r} / \hbar} / \sqrt{V}$ when the quantization is done within a large volume $V$. Thus

$$
V_{f i}=-\frac{e}{m_{e}} \sqrt{\frac{\hbar}{2 \omega \epsilon_{0} V^{2}}} \int d^{3} x e^{i \mathbf{p} \cdot \mathbf{r} / \hbar}\left(-i \hbar \mathbf{e}_{\lambda} \cdot \boldsymbol{\nabla}\right) e^{i \mathbf{k} \cdot \mathbf{r}} \psi_{i}(\mathbf{r})
$$

where $\psi_{i}(\mathbf{r})$ is the wave function for the bound electron. A partial integration now allows us to write this as

$$
\begin{equation*}
V_{f i}=\frac{e}{m_{e}} \sqrt{\frac{\hbar}{2 \omega \epsilon_{0} V^{2}}}\left(\mathbf{e}_{\lambda} \cdot \mathbf{p}\right) \tilde{\psi}_{i}(q) \tag{7.113}
\end{equation*}
$$

when we introduce the remaining integral

$$
\begin{equation*}
\tilde{\psi}_{i}(\mathbf{q})=\int d^{3} x \psi_{i}(\mathbf{r}) e^{-i \mathbf{q} \cdot \mathbf{r} / \hbar} \tag{7.114}
\end{equation*}
$$

where $\mathbf{q}=\mathbf{p}-\hbar \mathbf{k}$ is the momentum transfer to the final electron. We recoqnize the integral as the Fourier transform of the initial state wave function $\psi_{i}(r)$.

We now specialize to the simplest case with the initial electron in the ground state $\psi_{0}(r)=e^{-r / a} / \sqrt{\pi a^{3}}$ where $a$ is the Bohr radius as derived in Chapter 5. The integral (7.114) is then very similar to the one we encountered in Chapter 6 when we calculate the scattering amplitude for a Yukawa potential. In fact, we can now do the needed integral

$$
\begin{equation*}
\tilde{\psi}_{0}(q)=\sqrt{\frac{1}{\pi a^{3}}} \int d^{3} x e^{-r / a} e^{-i \mathbf{q} \cdot \mathbf{r} / \hbar} \tag{7.115}
\end{equation*}
$$

by writing it as

$$
\tilde{\psi}_{0}(q)=-4 \pi \sqrt{\frac{1}{\pi a^{3}}} \frac{\partial}{\partial \kappa} \int d^{3} x \frac{e^{-\kappa r}}{4 \pi r} e^{-i \mathbf{q} \cdot \mathbf{r} / \hbar}
$$

with $\kappa=1 / a$ and take over directly the screened Coulomb integral (6.57), giving

$$
\begin{align*}
\tilde{\psi}_{0}(q) & =-4 \pi \sqrt{\frac{1}{\pi a^{3}}} \frac{\partial}{\partial \kappa} \frac{1}{\kappa^{2}+q^{2} / \hbar^{2}}=\sqrt{\frac{1}{\pi a^{3}}} \frac{8 \pi \kappa}{\left(\kappa^{2}+q^{2} / \hbar^{2}\right)^{2}} \\
& =\sqrt{\frac{1}{\pi a^{3}}} \frac{8 \pi a^{3}}{\left(1+a^{2} q^{2} / \hbar^{2}\right)^{2}} \tag{7.116}
\end{align*}
$$

It depends on the squared momentum transfer $q^{2}=p^{2}+(\hbar k)^{2}-2 \hbar k p \cos \theta$ where $\theta$ is the angle between the photon momentum $\hbar \mathbf{k}$ and the electron momentum $\mathbf{p}$. With our kinematical constraints we can ignore the middle term and set $p^{2}=2 m_{e} \hbar \omega$. Thus we can take $q^{2}=2 m_{e} \hbar \omega(1-\beta \cos \theta)$ in the following. Since the Bohr radius $a=\hbar / Z \alpha m_{e} c$, it follows also that $a q / \hbar \gg 1$ from the same assumptions, and we can thus ignore the first term in the numerator of (7.116).
Using now the matrix element (7.113) in the Golden Rule, we find for the differential transition rate into a final state defined by the electron momentum $\mathbf{p}$ the result

$$
d \Gamma=\frac{2 \pi}{\hbar}\left(\frac{e}{m_{e}}\right)^{2} \frac{\hbar}{2 \omega \epsilon_{0} V^{2}}\left(\mathbf{e}_{\lambda} \cdot \mathbf{p}\right)^{2}\left|\tilde{\psi}_{0}(q)\right|^{2} \delta\left(E_{f}-\hbar \omega-E_{i}\right) \frac{V d^{3} p}{(2 \pi \hbar)^{3}}
$$

If the final electron is detected in the solid angle $d \Omega$, we can write $d^{3} p=p^{2} d p d \Omega$. Integrating now over the final momentum using $d E_{f}=p d p / m_{e}$, the $\delta$-function disappears. Using then the result (7.116) for $a q / \hbar \gg 1$, one is left with

$$
\frac{d \Gamma}{d \Omega}=\frac{2 \pi}{\hbar}\left(\frac{e}{m_{e}}\right)^{2} \frac{\hbar p m_{e}}{2 \omega \epsilon_{0} V}\left(\mathbf{e}_{\lambda} \cdot \mathbf{p}\right)^{2} \frac{8 \hbar^{5}}{\pi^{2} a^{5} q^{8}}
$$

We see that the electron prefers to be ejected along the polarization vector $\mathbf{e}_{\lambda}$, i.e. normal to the incoming photon direction $\mathbf{k}$. No electrons are produced exactly along this direction.

From here we obtain the differential cross-section by dividing by the incoming flux $I_{0}=c / V$ of photons. Assuming that they are unpolarized, we take the average by summing over the two possible polarization vectors using (7.67) and divide by two, resulting in

$$
\frac{d \sigma}{d \Omega}=\frac{2 \pi}{\hbar}\left(\frac{e}{m_{e}}\right)^{2} \frac{\hbar p m_{e}}{2 \omega \epsilon_{0} c}\left(\mathbf{p}^{2}-(\mathbf{p} \cdot \widehat{\mathbf{k}})^{2}\right) \frac{4 \hbar^{5}}{\pi^{2} a^{5} q^{8}}
$$

where $\mathbf{p}^{2}-(\mathbf{p} \cdot \widehat{\mathbf{k}})^{2}=p^{2} \sin ^{2} \theta$ when $\widehat{\mathbf{k}}$ is a unit vector in the direction of $\mathbf{k}$. Inserting here our approximation for $q^{2}$ and simplifying, we finally have for the differential cros-section

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=2 \sqrt{2} Z^{5} \alpha^{4} r_{0}^{2}\left(\frac{m_{e} c^{2}}{\hbar \omega}\right)^{7 / 2} \frac{\sin ^{2} \theta}{(1-\beta \cos \theta)^{4}} \tag{7.117}
\end{equation*}
$$

where $r_{0}=e^{2} / 4 \pi \epsilon_{0} m_{e} c^{2}$ is called the classical radius of the electron. Putting in numbers, it has the magnitude $r_{0}=2.82 \times 10^{-13} \mathrm{~cm}$ which for some strange reason is almost equal to the radius of a nucleus. The denominator in the result increases the photoproduction of electrons towards the forward direction. This effect is seen to become more pronounced at higher energies where the cross-section becomes smaller.

Since the cross-section varies like $Z^{5}$, it follows that the stronger the electron is bound to the nucleus, the easier it is to eject it. This seems somewhat paradoxical since one would naively think that the stronger the electron is bound, the more energy is needed to eject it. That is partially true, but at the same time the cross-section decreases with higher energy. Anyway, the net result is that photoproduction is most likely to take place for electrons in $K$ shell which is nearest to the nucleus. This also applies to more complex atoms than hydrogen we have considered here. Good agreement is then found when comparing with experiments.

### 7.13 Thomson scattering

When a photon hits a free electron, it cannot be absorbed and continue as a free electron. This follows from energy and momentum conservation. On the other hand, if the absorption is followed by re-emission of a new photon, we have an allowed process. Then we say that the photon has been scattered by the electron. We will now calculate the cross-section for this process.

Again we can use the basic interaction Hamiltonian (7.42) when we ignore the Pauli spin term in (7.30). There will be one photon in the initial state and also one in the final state. It will thus involve altogether two photons. Using perturbations theory, we can then calculate the scattering amplitude by considering the first term $\propto \mathbf{p} \cdot \mathbf{A}$ in (7.42) in a second-order calculation or the last term $\propto \mathbf{A}^{2}$ to first order in perturbation theory.

The electron causing the scattering, can in general be bound in an atom or free. The first case has very many important consequences, but is more elaborate to calculate. So we will here consider the simpler case of scattering on a free electron initially at rest, i.e. $\mathbf{p}_{i}=0$. Due to the interaction, it will receive a kick from the photon and recoil with a momentum $\mathbf{p}_{f}$ that follows from energy and momentum conservation. But if we now simplify even more and assume that the photon has an energy $\hbar \omega_{i}$ much less than the electron rest energy $m_{e} c^{2}$, then we see that the maximum energy $\left(\hbar \omega_{i}\right)^{2} / 2 m_{e} c^{2}$ which the kicked electron can carry away, will be much less than the incoming photon energy. The photon thus suffers no energy loss in the scattering process. Only the photon momentum will change direction, from the initial value $\hbar \mathbf{k}_{i}$ to the final value $\hbar \mathbf{k}_{f}$. This low-energy, elastic scattering of photons is called Thomson scattering and is important in many places. For instance, it is the driving meachanism in creating the characteristic spectrum for the cosmic microwave background temperature fluctuations.

Since we now can neglect the electron momentum, only the second term $\propto \mathbf{A}^{2}$ in (7.42) will contribute. And we only need to calculate its contribution in lowest order of perturbation theory. The initial state is then $|i\rangle=\left|\mathbf{k}_{i} \lambda_{i}\right\rangle$ when we include the photon polarization. It is detected in the corresponding final state $|f\rangle=\left|\mathbf{k}_{f} \lambda_{f}\right\rangle$ so that transi-
tion matrix element can be written as

$$
\begin{equation*}
V_{f i}=\frac{e^{2}}{2 m_{e}}\left\langle\mathbf{k}_{f} \lambda_{f}\right| \widehat{\mathbf{A}} \cdot \widehat{\mathbf{A}}\left|\mathbf{k}_{i} \lambda_{i}\right\rangle \tag{7.118}
\end{equation*}
$$

when we place the electron at position $\mathbf{x}=0$. Squaring the photon field operator (7.86), there will now two terms which contribute equally giving the simple result

$$
\begin{equation*}
V_{f i}=\frac{e^{2}}{m_{e}} \sqrt{\frac{\hbar}{2 \omega_{f} \epsilon_{0} V}} \sqrt{\frac{\hbar}{2 \omega_{i} \epsilon_{0} V}} \mathbf{e}_{\lambda^{\prime}}^{*} \cdot \mathbf{e}_{\lambda} \tag{7.119}
\end{equation*}
$$

This we can understand from considering the relevant matrix element which contains the factor $\left\langle\mathbf{k}_{f}\right| \widehat{a}_{\mathbf{k}} \widehat{a}_{\mathbf{k}^{\prime}}^{\dagger}+\widehat{a}_{\mathbf{k}}^{\dagger} \widehat{a}_{\mathbf{k}^{\prime}}\left|\mathbf{k}_{i}\right\rangle$ when we suppress the polarizations. It is obviously equal to $\langle 0| \widehat{a}_{\mathbf{k}_{f}} \widehat{a}_{\mathbf{k}} \widehat{a}_{\mathbf{k}^{\prime}}^{\dagger} \widehat{a}_{\mathbf{k}_{i}}^{\dagger}|0\rangle+\langle 0| \widehat{a}_{\mathbf{k}_{f}} \widehat{a}_{\mathbf{k}}^{\dagger} \widehat{a}_{\mathbf{k}^{\prime}} \widehat{a}_{\mathbf{k}_{i}}^{\dagger}|0\rangle$. In the first term we find one contribution from the commutator of $\widehat{a}_{\mathbf{k}_{f}}$ with $\widehat{a}_{\mathbf{k}^{\prime}}^{\dagger}$ while $\widehat{a}_{\mathbf{k}}$ is commuted with $\widehat{a}_{\mathbf{k}_{i}}^{\dagger}$ which gives $\delta_{\mathbf{k}_{f} \mathbf{k}^{\prime}} \delta_{\mathbf{k}_{i} \mathbf{k}}$. There is also a contribution coming from the commutator between $\widehat{a}_{\mathbf{k}_{f}}$ and $\widehat{a}_{\mathbf{k}_{i}}^{\dagger}$ and $\widehat{a}_{\mathbf{k}}$ commuted with $\widehat{a}_{\mathbf{k}^{\prime}}^{\dagger}$, thus giving $\delta_{\mathbf{k}_{f} \mathbf{k}_{i}} \delta_{\mathbf{k k}^{\prime}}$ and therefore only contributing when $\mathbf{k}_{i}=\mathbf{k}_{f}$. This implies no scattering, and we can ignore it. Similarly, from the second term we find a non-zero contribution resulting in $\delta_{\mathbf{k}_{f} \mathbf{k}} \delta_{\mathbf{k}_{i} \mathbf{k}^{\prime}}$. It is exactly the same as we just found from the first term, only with $\mathbf{k}$ and $\mathbf{k}^{\prime}$ interchanged. But since we sum over these wavevectors, both terms give exactly the same.

Using again Fermi's Golden Rule, we then have for the differential transition rate

$$
d \Gamma=\frac{2 \pi}{\hbar}\left(\frac{e^{2}}{m_{e}}\right)^{2}\left(\mathbf{e}_{f}^{*} \cdot \mathbf{e}_{i}\right)^{2}\left(\frac{\hbar^{2}}{2 \omega_{f} \epsilon_{0} V \cdot 2 \omega_{i} \epsilon_{0} V}\right) \delta\left(\hbar \omega_{f}-\hbar \omega_{i}\right) \frac{V d^{3} k_{f}}{(2 \pi)^{3}}
$$

in analogy with what we did with particle scattering in Chapter 6. If the initial photon


Figure 7.8: Scattering of light coming in along the $z$-axis and with initial polarization along $b f e_{i}$.
momentum is along the $z$-axis and the final photon is scattered an angle $\theta$ in the $x z$-plane as shown in Fig.7.8, then we have $d^{3} k_{f}=k_{f}^{2} d k_{f} d \Omega$ where $d \Omega=2 \pi \sin \theta d \theta$. Dividing by the incoming flux $I_{0}=c / V$ of photons and integrating over $\omega_{f}=c k_{f}$, the differential cross-section $d \sigma=d \Gamma / I_{0}$ becomes

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=r_{0}^{2}\left(\mathbf{e}_{f}^{*} \cdot \mathbf{e}_{i}\right)^{2} \tag{7.120}
\end{equation*}
$$

where $r_{0}=e^{2} / 4 \pi \epsilon_{0} m_{e} c^{2}$ again is the classical electron radius. Since the result is independent of the Planck-Dirac constant $\hbar$, it should also be derivable from classical electrodynamics which in fact J.J. Thomson did a hundred years ago after he had discovered the electron in 1898.

For a given direction of the incoming polarization vector $\mathbf{e}_{i}$, the Thomson result gives a differential cross-section which varies with the direction of the polarization vector of the scattered photon. If this final polarization is not observed, we must then sum over the two possible directions of $\mathbf{e}_{f}$ using the basic formula (7.67). It then follows that

$$
\frac{d \sigma}{d \Omega}=r_{0}^{2}\left[\left(\mathbf{e}_{i} \cdot \mathbf{e}_{i}\right)-\left(\mathbf{e}_{i} \cdot \hat{\mathbf{k}}_{f}\right)\left(\mathbf{e}_{i} \cdot \hat{\mathbf{k}}_{f}\right)\right]
$$

where $\hat{\mathbf{k}}_{f}$ is the unit vector along the final photon momentum $\hbar \mathbf{k}_{f}$. From the above figure we see that $\boldsymbol{\varepsilon}_{i} \cdot \hat{\mathbf{k}}_{f}=\cos \alpha=\sin \theta \cos \phi$ where $\phi$ is the angle in the $x y$-plane of the initial photon polarization. Thus we get

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=r_{0}^{2} \sin ^{2} \alpha \tag{7.121}
\end{equation*}
$$

for the differential cross-section. This is also the classical result for scattering of polarized light.

When the incoming photon is unpolarized, we must average over the two possible directions of the polarization vector $\mathbf{e}_{i}$. From (7.120) we then obtain the cross-section

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=\frac{r_{0}^{2}}{2}\left[\left(\mathbf{e}_{f} \cdot \mathbf{e}_{f}\right)-\left(\mathbf{e}_{f} \cdot \hat{\mathbf{k}}_{i}\right)\left(\mathbf{e}_{f} \cdot \hat{\mathbf{k}}_{i}\right)\right] \tag{7.122}
\end{equation*}
$$

The scattering plane formed by the photon momenta $\mathbf{k}_{i}$ and $\mathbf{k}_{f}$, is just the $x z$-plane with normal $\mathbf{n}=\mathbf{k}_{i} \wedge \mathbf{k}_{f}$ along the $y$-axis for our choice of coordinates. If we now choose the two directions of the polarization vector $\mathbf{e}_{f}$ to be in or normal to this plane, the polarized cross-sections becomes

$$
\frac{d \sigma}{d \Omega}=\frac{r_{0}^{2}}{2} \begin{cases}1 & \text { for } \quad \mathbf{e}_{f} \| \mathbf{n}  \tag{7.123}\\ 1-\sin ^{2} \theta & \text { for } \\ \mathbf{e}_{f} \perp \mathbf{n}\end{cases}
$$

In the special case when the photon is scattered by $\theta=90^{\circ}$, it is completely polarized normal to the scattering plane.

When we don't observe the polarization of the final photon, we must sum over these two contributions and find

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=\frac{r_{0}^{2}}{2}\left(1+\cos ^{2} \theta\right) \tag{7.124}
\end{equation*}
$$

It has a minimum for $\theta=90^{\circ}$ and maxima in the forward and backward directions.
The above unpolarized cross-section can also be obtained directly from the master formula (7.67) when summing over $\mathbf{e}_{f}$ in the result (7.122). We then obtain

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=\frac{r_{0}^{2}}{2}\left[2-\left(\hat{\mathbf{k}}_{i} \cdot \hat{\mathbf{k}}_{i}\right)+\left(\hat{\mathbf{k}}_{i} \cdot \hat{\mathbf{k}}_{f}\right)^{2}\right] \tag{7.125}
\end{equation*}
$$

which gives the same result since $\hat{\mathbf{k}}_{i} \cdot \hat{\mathbf{k}}_{f}=\cos \theta$. The total cross-section is now found by integrating over all directions of the scattered photon. It becomes

$$
\begin{equation*}
\sigma=\pi r_{0}^{2} \int_{-1}^{1} d \cos \theta\left(1+\cos ^{2} \theta\right)=\frac{8}{3} \pi r_{0}^{2} \tag{7.126}
\end{equation*}
$$

which is the total Thomson cross-section. Had we included the non-relativistic corrections of the order $\hbar \omega_{i} / m c^{2}$, it would have been multiplied by the factor $1-2 \hbar \omega_{i} / m c^{2}$. The cross-section thus decreases with increasing energy. At higher energies this process is called Compton scattering and must be calculated using relativistic quantum mechanics for the electron.

As already advertised, the above Thomson cross-section can also be derived from classical electrodynamics. Then one consideres an incoming light wave along the $z$-axis with an electric filed $\mathbf{E}_{i}$ in the $x y$-plane making an angle $\phi$ with the $x$-axis. It gives an acceleration $\dot{\mathbf{v}}=e \mathbf{E}_{i} / m_{e}$ to the electron. This again generates a new electric field which in the radiation zone is

$$
\begin{equation*}
\mathbf{E}_{f}(\mathbf{r})=\frac{e}{4 \pi \epsilon_{0} c^{2} r} \hat{\mathbf{r}} \wedge(\hat{\mathbf{r}} \wedge \dot{\mathbf{v}}) \tag{7.127}
\end{equation*}
$$

where $\hat{\mathbf{r}}$ is a unit vector in the direction of the field point $\mathbf{r}$. The electric field is accompanied by the corresponding magnetic field $H_{f}=\sqrt{\epsilon_{0} / \mu_{0}} E_{f}$ and is radiated away. The resulting Poynting vector $\mathbf{S}=\mathbf{E}_{f} \wedge \mathbf{H}_{f}$ has then the magnitude

$$
|\mathbf{S}|=I_{0}\left(\frac{r_{0}}{r}\right)^{2} \sin ^{2} \alpha
$$

where $I_{0}=\sqrt{\epsilon_{0} / \mu_{0}} E_{i}^{2}$ is the incoming energy flux. In this classical calculation the angle $\alpha$ is now between the initial electric vector $\mathbf{E}_{i}$ and the field point defined by $\mathbf{r}$. In the quantum description illustrated in Fig.7.8 there is no final state field point, but instead the wave vector $\mathbf{k}_{f}$ of the outgoing photon. From the Poynting vector one now finds the scattered light intensity $d I=r^{2}|\mathbf{S}| d \Omega$ into the solid angle $d \Omega$ to be $d I=I_{0} r_{0}^{2} \sin ^{2} \alpha d \Omega$. The differential cross-section $d \sigma=d I / I_{0}$ then follows as

$$
\begin{equation*}
\frac{d \sigma}{d \Omega}=r_{0}^{2} \sin ^{2} \alpha \tag{7.128}
\end{equation*}
$$

in agreement with (7.121) when the incoming light is polarized. Writing now $\sin ^{2} \alpha=$ $1-\cos ^{2} \phi \sin ^{2} \theta$ and averaging over this polarization with $\left\langle\cos ^{2} \phi\right\rangle=1 / 2$, one then easily finds the result (7.124) for unpolarized light.

### 7.14 Vaccum energies and the cosmological constant

We have already seen that the quantization of the electromagnetic field leads to a divergent value for the vacuum energy. This is a very surprising and disturbing result. Not so much perhaps when we don't include the effects of gravitation since then we just have to consider energy differences where such an infinite contribution would cancel out. But including gravitation and Einstein's equivalence between energy and mass, it
would have disastrous consequences. This is simplest to discuss within cosmology where Einstein introduced the so-called cosmological constant $\Lambda$. In the last few it has been measured by astronomers using different different methods and they have found a value close to $\Lambda=10^{-54} \mathrm{~cm}^{-2}$. It results in an acceleration of the expansion of the Universe. Had this constant been much bigger, the acceleration would have been much larger and we would not have been here.

The cosmological constant corresponds to the vacuum energy density $\mathcal{E}_{\Lambda}=\Lambda c^{4} / 8 \pi G$ of the Universe. Here $G$ is Newton's gravitational constant. We don't know where this energy is coming from and it is often called dark energy. With the above value for $\Lambda$ we find that $\mathcal{E}_{\Lambda}=10^{-29} \mathrm{gc}^{2} / \mathrm{cm}^{3}$ which is equivalent to the energy of two hydrogen atoms per cubic meter. It's very close to the density of real matter in the Universe. But this is not made up of ordinary atoms, but of so-called dark matter which we also don't know anything about. But that is a different story.

Let us now compare this with the contribution from the vaccum fluctuations of the electromagnetic field. Considering the energy in a very large volume $V$ so that we can replace the summation in the expression (7.83) by an integration and including the contributions from the two polarization directions, we have

$$
\begin{equation*}
E_{0}=V \hbar c \int \frac{d^{3} k}{(2 \pi)^{3}} k=V \frac{\hbar c}{2 \pi^{2}} \int_{0}^{\infty} d k k^{3} \tag{7.129}
\end{equation*}
$$

since the frequency $\omega=c k$. This integral certainly gives an infinity in the upper limit where $k \rightarrow \infty$, i.e. at very small distances. The simplest way to make it finite is to accept that our theory is not valid up to arbitrary high momenta, but only up to wave number $K$. This new quantity will act as a cut-off in the integration. Physically it means that the present quantum theory is only valid down to distances of size $L=1 / K$. For the equivalent vacuum energy density $\mathcal{E}_{0}=E_{0} / V$ we then have

$$
\begin{equation*}
\mathcal{E}_{0}=\frac{\hbar c}{2 \pi^{2}} \int_{0}^{K} d k k^{3}=\frac{\hbar c}{8 \pi^{2}} K^{4} \tag{7.130}
\end{equation*}
$$

So what can the size of this new quantity $K$ be? In high energy physics experiments up to several hundred GeV we have seen no trace of such a cut-off. Thus $K>10^{-16} \mathrm{~cm}$. Probably it is much smaller. We thus get the limit

$$
\mathcal{E}_{0}>\frac{\hbar c}{8 \pi^{2}} \times 10^{64} \mathrm{~cm}^{-4}
$$

for the quantum-induced vacuum energy. Since $\hbar c=197 \mathrm{eV} \times 10^{-7} \mathrm{~cm}$ we thus have the limit $\mathcal{E}_{0}>10^{57} \mathrm{eV} / \mathrm{cm}^{3}$. To compare it with the cosmological, dark energy, we make use of the conversion factor $1 \mathrm{eV} / \mathrm{c}^{2}=1.783 \times 10^{-33} \mathrm{~g}$ which gives $\mathcal{E}_{0}>10^{24} \mathrm{gc}^{2} / \mathrm{cm}^{3}$. We then have then have the chocking result that it is larger than the observed vacuum energy density by a factor $10^{53}$ !! So here there is something fundamentally wrong with the theory and nobody knows why. This is probably the most important, unsolved problem in theoretical physics today.
We would like to iknow if there is some upper theoretical limit the size of the cut-off $K$ ? A quantum fluctuation on this scale will have a corresponding momentum $P=\hbar K$ and therefore also energy $E=c P=\hbar c K$. This can be considered to be a direct consequence
of the Heisenberg uncertainty principle. For the resulting mass $M=E / c^{2}=\hbar K / c$ to manifest itself and not form a black hole, its Compton wavelength $\hbar / M c$ must be larger than the size of the black hole it can form, i.e. $\hbar / M c>G M / c^{2}$. We thus have the limit $M>M_{P}$. Here $M_{P}=\sqrt{\hbar c / G}=10^{19} \mathrm{GeV} / \mathrm{c}^{2}$ is called the Planck mass after the same Max Planck we know from the theory of blackbody radiation and who also was the first to dicover this limiting mass. This now corresponds to a smaller length $L_{P}=1 / K_{P}=\sqrt{\hbar G / c^{3}}=10^{-33} \mathrm{~cm}$. Fluctuations with smaller wavelengths will be so massive that they will form black holes and our whole concept of a classical spacetime breaks down on these short scales.

Assumung now that standard physics is valid down this small Planck length, we must use the corresponding cut-off $K_{P}$ in the expression (7.130) for the quantum-induced vacuum energy. The result is obviously even bigger than before. In fact, the discrepancy with the observed value is now seen to be factor $10^{53} \times 10^{17 \cdot 4}=10^{121}$. Never in the history of physics have there been a result so totally wrong!!! This is one reason why many hope that new physics like string theory will be needed to solve this conundrum. In many ways the situation is very similar to the problems with blackbody radiation a hundred years ago before Planck postulated that the energy of a system could not change continuously, but in discrete quanta. Thus quantum mechanics was born. Today we probably need a similar revolution in our understanding of the world.

Apparently the problem lies in the quantum result (7.83) for the vacuum energy. One immediate way out of the conundrum is to conclude that there is something wrong with this expression. From the way it appeared in our calculation, it's origin can be traced back to the zero-point energy of the harmonic oscillator. This is equivalent to the nonzero, canonical commutator of the position and conjugate momentum operators. And to abondon this, would be a total collapse of what we know as quantum mechanics.

### 7.15 The Casimir force

But there is good physics in the expression (7.83) for the quantum-induced vacuum energy. This is the Casimir effect. Consider two parallel plates with separation $L$. Take their common normal $\mathbf{n}$ be along the $z$-axis as shown in the figure. Assume that they are made by metal so that whatever electric and magnetic fields are present, they must satisfy the corresponding boundary conditions $\mathbf{n} \times \mathbf{E}=0$ and $\mathbf{n} \cdot \mathbf{B}=0$ which follow from standard electromagnetic theory. These force the plane wave modes (7.60) to be replaced by more complicated functions. In order to simplify the calculation, we assume instead that the vector potential A must be zero at the plates. We then make a small mistake, which we will came back to in a moment. The classical modes will thus contain a standing wave in the $z$-direction, i.e.

$$
\begin{equation*}
\mathbf{A}_{n \mathbf{k}_{T}}(\mathbf{x}, t)=\sin (n z \pi / L) e^{i\left(\mathbf{k}_{T} \cdot \mathbf{x}_{T}-\omega t\right)} \tag{7.131}
\end{equation*}
$$

before we normalize them. Here $\mathbf{k}_{T}=\left(k_{x}, k_{y}\right)$ is the wave number parallel to the $x y$ plane while $\mathbf{x}_{T}=(x, y)$ is the position vector in the same plane. The wave number along the $z$-axis is $k_{z}=n \pi / L$ where the quantum number $n=1,2,3, \ldots$. For such a mode the frequency is therefore $\omega=c\left(\mathbf{k}_{T}^{2}+k_{z}^{2}\right)^{1 / 2}$. The resulting cacuum energy can


Figure 7.9: Two parallel plates are attracted to each other by the Casimir force.
then be written as

$$
\begin{equation*}
E_{0}=A \hbar \sum_{n=1}^{\infty} \int \frac{d^{2} k_{T}}{(2 \pi)^{2}} \sqrt{\mathbf{k}_{T}^{2}+(n \pi / L)^{2}} \tag{7.132}
\end{equation*}
$$

as follows from (7.83) and (7.129). Here $A$ is the cross-sectional area of the plates in the limit where it becomes very large. We have also included a factor 2 from the two polarization directions. Again we see a divergence both in the integration and the summation. In order to understand this better, we introduce $\omega d \omega=c^{2} k_{T} d k_{T}$ in integral so that it becomes

$$
E_{0}=\frac{A}{2 \pi c^{2}} \sum_{n=1}^{\infty} \int_{\kappa c n}^{\infty} d \omega \omega^{2}
$$

where $\kappa=\pi / L$. We can now dampen the divergence in the upper limit where $\omega \rightarrow \infty$ by multiplying the integrand by an exponential regulator,

$$
E_{0}^{r e g}=\frac{A \hbar}{2 \pi c^{2}} \sum_{n=1}^{\infty} \int_{\kappa c n}^{\infty} d \omega \omega^{2} e^{-\omega \varepsilon / c}
$$

Here $\varepsilon$ is very small length. As long as it is finite, the integral is also finite. At the end of the calculation we must then consider the limit $\varepsilon \rightarrow 0$. The integral can now be evaluated by standard methods, giving

$$
\begin{align*}
E_{0}^{r e g} & =\frac{A \hbar}{2 \pi} \sum_{n=1}^{\infty} \frac{\partial^{2}}{\partial \varepsilon^{2}} \int_{\kappa c n}^{\infty} d \omega e^{-\omega \varepsilon / c} \\
& =\frac{A \hbar}{2 \pi} \frac{\partial^{2}}{\partial \varepsilon^{2}} \frac{c}{\varepsilon} \sum_{n=1}^{\infty} e^{-\kappa n \varepsilon}=\frac{A \hbar c}{2 \pi} \frac{\partial^{2}}{\partial \varepsilon^{2}} \frac{1}{\varepsilon} \frac{1}{e^{\kappa \varepsilon}-1} \tag{7.133}
\end{align*}
$$

Since we are eventually interested in the limit $\varepsilon \rightarrow 0$, we can here use the expansion

$$
\begin{equation*}
\frac{x}{e^{x}-1}=\sum_{n=0}^{\infty} \frac{B_{n}}{n!} x^{n}=1-\frac{x}{2}+\frac{x^{2}}{12}-\frac{x^{4}}{720}+\cdots \tag{7.134}
\end{equation*}
$$

based on the Bernoulli numbers $B_{n}$. The lowest ones are $B_{0}=1, B_{1}=-1 / 2, B_{2}=1 / 6$ and $B_{4}=-1 / 30$. It is now straightforward to take the double derivative and calculate the regulated vacuum energy. For the energy density $\mathcal{E}_{0}^{\text {reg }}=E_{0}^{\text {reg }} / A L$ we then obtain the result

$$
\begin{equation*}
\mathcal{E}_{0}^{r e g}=\frac{3 \hbar c}{\pi^{2} \varepsilon^{4}}-\frac{\hbar c}{2 \pi \varepsilon^{3} L}-\frac{\pi^{2} \hbar c}{720 L^{4}} \tag{7.135}
\end{equation*}
$$

when we neglect all term involving positive powers of $\varepsilon$.
We are seen to still have a divergent in vacuum energy, but it is now isolated in the two first terms. Having no plates in the vacuum corresponds to taking them away from each other so that $L \rightarrow \infty$. Then we are left with just the first term. It is seen to correspond to the previous vacuum energy (7.130) where the wave number cut-off is now replaced by $K \propto 1 / \varepsilon$. The two next terms in (7.135) thus correspond to a change in the vacuum energy induced by the presence of the plates. Of these the first is still divergent in the limit $\varepsilon \rightarrow 0$ and could be worrisome. But a more careful calculation taking into account the correct boundary conditions mentioned above, shows that this term will then actually be absent. Thus, the net effect of the two plates is just the last term in (7.135) and is called the Casimir energy density,

$$
\begin{equation*}
\mathcal{E}_{C}=-\frac{\pi^{2} \hbar c}{720 L^{4}} \tag{7.136}
\end{equation*}
$$

The closer the plates are to each other, the more the energy decreases. Since this system strives towards an energy minimum, the two plates are therefore attracted towards each other by what is now called the Casimir force after H. Casimir who derived this result in 1948. Its size $F$ is given by the derivative of the energy $A L \mathcal{E}_{C}$ between the plates with respect to the plate separation $L$. Per unit area it corresponds to the pressure $P=F / A$ which therefore becomes

$$
\begin{equation*}
P=-\frac{\pi^{2} \hbar c}{240 L^{4}} \tag{7.137}
\end{equation*}
$$

Putting in numbers it gives

$$
\begin{equation*}
P=-40\left(\frac{100 \mathrm{~nm}}{L}\right)^{4} \mathrm{~N} / \mathrm{m}^{2} \tag{7.138}
\end{equation*}
$$

The effect was first experimentally verified by M. Sparnay in 1958 and confirmed with much greater accuracy during the last few years. So there is no longer any doubt that the quantum-induced vacuum energy exists, but its relation to the cosmologically observed value is still a big mystery.

## Chapter 8

## Spin interactions

In chapter 5 it was shown how the bound motion of a non-relativistic, charged particle in a Coulomb potential could be obtained exactly in quantum mechanics. This solution provides the framework for all atomic physics, both in hydrogen when higher order effects are included and in atomes with more than one electron. These higher order effects are of two main types, both of basically relativistic origin. There will be a higher order kinematic correction because the electron moves with a velocity close to one percent of the velocity of light. In addition, the motion of both the charged electron and nucleus sets up magnetic fields which will couple to the spins of these particles. These effects must be calculated in perturbation theory. Degeneracies forces us two find new basis states for the atom which are eigenstates of the total electron angular momentum. Rules for addition of spins are derived and the new states defines the corresponding ClebschGordan cofficients. A more realistic calculation of the Zeeman effect and the magnetic moment of the atom can then be obtained.

### 8.1 Relativistic velocity correction

The energy of a relativistic particle with mass $m$ and momentum $p$ is according to Einstein $E=\left(p^{2} c^{2}+m^{2} c^{4}\right)^{1 / 2}$. When $p<m c$ it thus gives by the binomial expansion

$$
\begin{align*}
E & =m c^{2}\left[1+\frac{1}{2}\left(\frac{p}{m c}\right)^{2}-\frac{1}{8}\left(\frac{p}{m c}\right)^{4}+\cdots\right] \\
& =m c^{2}+\frac{p^{2}}{2 m}-\frac{p^{4}}{8 m^{3} c^{2}}+\cdots \tag{8.1}
\end{align*}
$$

While the rest energy in the first term usually is of no interest in atomic physics, the second term is the standard kinetic energy while the third term is the lowest, kinematic correction to the energy. For the electron bound in a hydrogen atom, it thus gives the perturbation

$$
\begin{equation*}
\widehat{H}_{v e l}=-\frac{1}{8 m_{e}^{3} c^{2}}(\widehat{\mathbf{p}} \cdot \widehat{\mathbf{p}})^{2} \tag{8.2}
\end{equation*}
$$

Since the velocity of the electron is close to one percent of the speed of light, it will give a significant shift of the energy levels in hydrogen calculated in Chapter 5. Fortunately
it is seen to have diagonal matrix elements in the non-perturbed basis so that these shifts are directly obtainable from

$$
\begin{equation*}
E_{v e l}=\left\langle n l m_{l}\right| \widehat{H}_{v e l}\left|n l m_{l}\right\rangle=-\frac{1}{2 m_{e} c^{2}}\left\langle n l m_{l}\right|\left(\frac{\widehat{\mathbf{p}}^{2}}{2 m_{e}}\right)^{2}\left|n l m_{l}\right\rangle \tag{8.3}
\end{equation*}
$$

since the spin parts of the full state vectors $\left|n l m_{l} ; m_{s}\right\rangle$ just factors out to give one. Now we can express the kinetic operator $\widehat{\mathbf{p}}^{2} / 2 m_{e}$ by the unperturbed atomic Hamiltonian $\widehat{H}$ so that

$$
E_{v e l}=-\frac{1}{2 m_{e} c^{2}}\left\langle n l m_{l}\right|\left(\widehat{H}+\frac{Z e^{2}}{4 \pi \epsilon_{0} \widehat{r}}\right)^{2}\left|n l m_{l}\right\rangle
$$

Since the external states in the matrix element are eigenstates of the Hamiltonian with eigenvalues which can be written as $E_{n}=-Z e^{2} / 8 \pi \epsilon_{0} a$ when expressed by the effective Bohr radius $a=a_{0} / Z$, we immediately have the result

$$
\begin{aligned}
E_{\text {vel }} & =-\frac{1}{2 m_{e} c^{2}}\left(\frac{Z e^{2}}{4 \pi \epsilon_{0}}\right)^{2}\left\langle\left(\frac{1}{r}-\frac{1}{2 a n^{2}}\right)^{2}\right\rangle \\
& =-\frac{1}{2 m_{e} c^{2}}\left(\frac{Z e^{2}}{4 \pi \epsilon_{0} a}\right)^{2}\left[\left\langle\frac{a^{2}}{r^{2}}\right\rangle-\frac{1}{n^{2}}\left\langle\frac{a}{r}\right\rangle+\frac{1}{4 n^{4}}\right]
\end{aligned}
$$

The radial expectation values here and in the following are then defined by

$$
\begin{equation*}
\left\langle\frac{1}{r^{s}}\right\rangle=\int d^{3} x\left|R_{n l}\right|^{2} r^{-s} \tag{8.4}
\end{equation*}
$$

and must be calculated for $s=1$ and $s=2$.
Since we know the exact wave functions, such expectation values can be obtained by direct integration for each particular energy level. To find a general formula valid for all the states, is more difficult. A first step is to calculate $\langle a / r\rangle$ which is like the expectation value of the Coulomb potential $V=-Z e^{2} / 4 \pi \epsilon_{0} r$. This can be obtained from the quantum mechanical virial theorem. It follows from the observation that for any operator $\widehat{A}$ without any explicit time dependence, the expectation value of the time derivative $d \widehat{A} / d t$ is zero. This follows from the Heisenberg equation of motion which gives for a general energy eigenstate $|n\rangle$

$$
\langle n| \frac{d \widehat{A}}{d t}|n\rangle=\frac{i}{\hbar}\langle n|[\widehat{H}, \widehat{A}]|n\rangle=\frac{i}{\hbar} E_{n}\langle n| \widehat{A}-\widehat{A}|n\rangle=0
$$

Now with $A=\mathbf{x} \cdot \mathbf{p}$ in the coordinate representation it follows

$$
\frac{d A}{d t}=\dot{\mathbf{x}} \cdot \mathbf{p}+\mathbf{x} \cdot \dot{\mathbf{p}}=\frac{1}{m} \mathbf{p}^{2}-\mathbf{x} \cdot \nabla V
$$

Thus we have the virial result $\left\langle\mathbf{p}^{2}\right\rangle=m\langle\mathbf{x} \cdot \boldsymbol{\nabla} V\rangle$. If the spherical potential varies like $V \propto r^{k}$, it gives the relation

$$
\begin{equation*}
\langle K\rangle=\frac{k}{2}\langle V\rangle \tag{8.5}
\end{equation*}
$$

between the expectation values of the kinetic energy $K=\mathbf{p}^{2} / 2 m$ and the potential energy $V$. In this form the virial theorem also exists in classical mechanics.

For the Coulomb potential $k=-1$ and therefore $\langle K\rangle=-\langle V\rangle / 2$ as we already have observed to be valid for the ground state. Since the full energy $E_{n}=-Z e^{2} / 8 \pi \epsilon_{0} a n^{2}$ now equals $E_{n}=\langle V\rangle / 2$ in general, we get the wanted expectation value

$$
\begin{equation*}
\left\langle\frac{a}{r}\right\rangle=\frac{1}{n^{2}} \tag{8.6}
\end{equation*}
$$

Combining this result with the radial Schrödinger equation, one can calculate all other similar expectation values. We will here need

$$
\begin{equation*}
\left\langle\frac{a^{2}}{r^{2}}\right\rangle=\frac{1}{n^{3}(l+1 / 2)} \tag{8.7}
\end{equation*}
$$

which also can be checked for the lowest levels by using the corresponding wave functions. Using now these results in the above relativistic correction, we obtain the final result

$$
\begin{align*}
E_{v e l} & =-\frac{1}{2 m_{e} c^{2}}\left(\frac{Z e^{2}}{4 \pi \epsilon_{0} a}\right)^{2} \frac{1}{n^{4}}\left(\frac{2 n}{2 l+1}-\frac{3}{4}\right) \\
& =-m_{e} c^{2}\left(\frac{Z \alpha}{n}\right)^{4}\left(\frac{n}{2 l+1}-\frac{3}{8}\right) \tag{8.8}
\end{align*}
$$

Since the last term is just a constant and therefore the same for all levels, the splitting between them is to be found in the first term. It lifts all levels slightly above the unperturbed levels of typical magnitudes set by $m_{e} c^{2}(Z \alpha)^{2}$. Relative to these energies, the relativistic shifts are therefore of the order of $(Z \alpha)^{2}$ smaller which is factor $\approx 10^{-5}$ in hydrogen. From the previous chapter we know that this corresponds to the effect of a magnetic field of typical size 1T acting on the electron spin. Accepting this point of view, it is a pretty big effect.

### 8.2 Spin-orbit splitting

In the rest frame of the nucleus its positive charge $Z e$ sets up the electric field

$$
\begin{equation*}
\mathbf{E}=\frac{Z e \mathbf{r}}{4 \pi \epsilon_{0} r^{3}} \tag{8.9}
\end{equation*}
$$

which binds the negatively charged electron moving with velocity $\mathbf{v}$. This is also present with the same magnitude in the rest frame of the electron. But here the nucleus is seen to move with velocity $\mathbf{- v}$ creating an additional small magnetic field given by the BiotSavart formula

$$
\begin{equation*}
\mathbf{B}=-\frac{\mu_{0}}{4 \pi} \frac{Z e \mathbf{v} \wedge \mathbf{r}}{r^{3}}=-\frac{1}{c^{2}} \mathbf{v} \wedge \mathbf{E} \tag{8.10}
\end{equation*}
$$

when introducing the velocity of light $c^{2}=1 / \epsilon_{0} \mu_{0}$. It couples to the spin $\mathbf{S}$ of the electron by the standard dipole coupling

$$
H_{e l}=-\boldsymbol{\mu} \cdot \mathbf{B}=\frac{e g_{e}}{2 m_{e}} \mathbf{S} \cdot \mathbf{B}
$$

From the beginning of the chapter we know that it causes the spin to precess around the field with the angular velocity $\omega=e g_{e} B / 2 m_{e}$ in this rest frame. But we want to describe the motion of the electron in the rest frame of the nucleus. Then there comes in an extra and purely relativistic effect called the Thomas precession, which reduces the overall precessional rate by an amount corresponding to the replacement $g_{e} \rightarrow g_{e}-1$ in the above formulas. Since we can here take $g_{e}=2$, this corresponds to half the original value for the precessional rate. We thus find the resulting interaction energy

$$
\begin{equation*}
H_{L S}=\frac{e}{2 m_{e}} \mathbf{S} \cdot \mathbf{B}=\frac{Z e^{2}}{4 \pi \epsilon_{0} r^{3}} \frac{\mathbf{L} \cdot \mathbf{S}}{2 m_{e}^{2} c^{2}} \tag{8.11}
\end{equation*}
$$

when we insert the electric field (8.9) into the magnetic field (8.10) and express the result in terms of the angular momentum $\mathbf{L}=\mathbf{r} \wedge m_{e} \mathbf{v}$ of the electron. This is called for obvious reasons the spin-orbit coupling. One might perhaps be sceptical to this semiclassical derivation, but it follows exactly from a fully relativistic derivation from the Dirac theory with the effect of the Thomas precession automatically taken care of.
The spin-orbit Hamiltonian contains a radial part times the part $\widehat{H}=2 \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}$ coupling the orbital spin $\widehat{\mathbf{L}}$ to the intrinsic spin $\widehat{\mathbf{S}}$. It will later turn out to be convenient to keep the factor 2 here. These operators act on the angular momentum product states $\left|l m_{l}\right\rangle\left|s m_{s}\right\rangle$ with $s=1 / 2$. We write them for short as $\left|m_{l} ; m_{s}\right\rangle$ where the $Z$-component of the orbital spin is given by $m_{l}=(l, l-1, \ldots,-l)$ and $m_{s}= \pm 1 / 2$. These can be taken as basis states in the corresponding $2(2 l+1)$-dimensional spin Hilbert space. In order to see the action of the Hamiltonian, we write it as

$$
\begin{equation*}
\widehat{H}=2 \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}=2 \widehat{L}_{z} \widehat{S}_{z}+\widehat{L}_{+} \widehat{S}_{-}+\widehat{L}_{-} \widehat{S}_{+} \tag{8.12}
\end{equation*}
$$

It will obviously not be diagonal in this basis because of the raising and lowering spin operators present here. Since the $2(2 l+1)$ states within the energy level specified by $\left|m_{l}, m_{s}\right\rangle$ have the same unperturbed energy, we must use degenerate perturbation theory to calculate the effect of the spin-orbit interaction. In practice, that means calculating the matrix of the Hamiltonian (8.12) in this basis and then diagonalize it.

Almost always it is simplest to start with the state with highest $z$-component of the angular momentum. That was the case with quantization of a single spin in Chapter 5 spin and is also true here. This state is now $|l ;+1 / 2\rangle$. Since neither spin can be raised higher here, it is an eigenspinor of the Hamiltonian with eigenvalue $l \hbar^{2}$,

$$
\begin{equation*}
\widehat{H}|l ;+1 / 2\rangle=\hbar^{2}[2 l(+1 / 2)+0+0]|l ;+1 / 2\rangle=l \hbar^{2}|l ;+1 / 2\rangle \tag{8.13}
\end{equation*}
$$

However, when applying this operator to a lower state with $m_{l}<l$, we obtain

$$
\begin{aligned}
\widehat{H}\left|m_{l} ;+1 / 2\right\rangle & =\hbar^{2} m_{l}\left|m_{l} ;+1 / 2\right\rangle \\
& +\hbar^{2} \sqrt{\left(l+1+m_{l}\right)\left(l-m_{l}\right)}\left|m_{l}+1 ;-1 / 2\right\rangle
\end{aligned}
$$

using the master formula (5.70) for the action of the spin ladder operators. The state $\left|m_{l} ;+1 / 2\right\rangle$ is seen to mix in with the state $\left|m_{l}+1 ;-1 / 2\right\rangle$. We must therefore apply the Hamiltonian also to this new state with the result

$$
\begin{aligned}
\widehat{H}\left|m_{l}+1 ;-1 / 2\right\rangle & =-\hbar^{2}\left(m_{l}+1\right)\left|m_{l}+1 ;-1 / 2\right\rangle \\
& +\hbar^{2} \sqrt{\left(l+1+m_{l}\right)\left(l-m_{l}\right)}\left|m_{l} ;+1 / 2\right\rangle
\end{aligned}
$$

No additional states are generated. The Hamiltonian can therefore be represented as a $2 \times 2$ matrix in this sub-space spanned by $\left|m_{l} ;+1 / 2\right\rangle$ and $\left|m_{l}+1 ;-1 / 2\right\rangle$. We can read off the matrix elements from the last two equations and the corresponding eigenvalues $\lambda \hbar^{2}$ are then seen to follow from

$$
\left|\begin{array}{cc}
m_{l}-\lambda & \sqrt{\left(l+1+m_{l}\right)\left(l-m_{l}\right)}  \tag{8.14}\\
\sqrt{\left(l+1+m_{l}\right)\left(l-m_{l}\right)} & -\left(m_{l}+1\right)-\lambda
\end{array}\right|=0
$$

It looks a bit complicated, but reduces to simply $\lambda(\lambda+1)=l(l+1)$. There is no dependence on the azimuthal quantum number $m_{l}$ in the two eigenvalues $\lambda_{+}=l$ and $\lambda_{-}=-(l+1)$ since the Hamiltonian (8.12) is rotational invariant. We notice that the positive eigenvalue is the same as for the highest state in (8.13).
Having the eigenvalues, the corresponding eigenstates are straightforward to find. They must be of the form $a\left|m_{l} ;+1 / 2\right\rangle+b\left|m_{l}+1 ;-1 / 2\right\rangle$ with $|a|^{2}+|b|^{2}=1$ when they are normalized. With the eigenvalue $\lambda_{+}=l$ we then find the relation

$$
\frac{a}{b}=\sqrt{\frac{l+1+m_{l}}{l-m_{l}}}
$$

between these two coefficients. The corresponding normalized eigenstate is therefore

$$
\begin{equation*}
\left|\lambda_{+}\right\rangle=\sqrt{\frac{l+1+m_{l}}{2 l+1}}\left|m_{l} ;+1 / 2\right\rangle+\sqrt{\frac{l-m_{l}}{2 l+1}}\left|m_{l}+1 ;-1 / 2\right\rangle \tag{8.15}
\end{equation*}
$$

when we choose all phase factors equal to one. For the other eigenvalue we can repeat the calculation or simply use that this new eigenstate must be orthogonal to the first. Then we have immediately

$$
\begin{equation*}
\left|\lambda_{-}\right\rangle=-\sqrt{\frac{l-m_{l}}{2 l+1}}\left|m_{l} ;+1 / 2\right\rangle+\sqrt{\frac{l+1+m_{l}}{2 l+1}}\left|m_{l}+1 ;-1 / 2\right\rangle \tag{8.16}
\end{equation*}
$$

There are $2 l$ of these states while there are $2 l+2$ eigenstates with the higher eigenvalue $\lambda_{+}$. The extra two comes from the previously considered highest state $|l ;+1 / 2\rangle$ and corresponding lowest state $|-l ;-1 / 2\rangle$ with the same eigenvalue $\lambda_{+}=l$. We thus see


Figure 8.1: Spin-orbit splitting of energy level in hydrogen.
that the total number of states $4 l+2=2(2 l+1)$ remains constant under this split as it should. What happens to one of the original unperturbed energy levels with $l \geq 1$ under the action of this spin-orbit perturbation is illustrated in Fig.8.1. Before we finish a more quantitative evaluation of its absolute magnitude in the hydrogen atom, we will develop an easier, more systematic and physical motivated method in treating systems with coupled spins.

### 8.3 Addition of spins

Degenerate perturbation theory gives us two ways to calculate the first order energy shift. We can choose an arbitrary basis and then diagonalize the matrix of the perturbation. This is the approach we have just completed. Or we can find a particular basis in which the perturbation will be diagonal. As we will now see, this method is simpler and gives a better physical understanding of the obtained results.
It is based on constructing the operator $\widehat{\mathbf{J}}=\widehat{\mathbf{L}}+\widehat{\mathbf{S}}$ for the total angular momentum of the electron and finding its eigenstates. It is almost trivial to see that it has the correct commutator to be a new spin operator. Indeed,

$$
\begin{aligned}
{\left[\widehat{J}_{i}, \widehat{J}_{j}\right] } & =\left[\widehat{L}_{i}, \widehat{L}_{j}\right]+\left[\widehat{S}_{i}, \widehat{S}_{j}\right]=i \hbar \epsilon_{i j k}\left(\widehat{L}_{k}+\widehat{S}_{k}\right) \\
& =i \hbar \epsilon_{i j k} \widehat{J}_{k}
\end{aligned}
$$

since the components of orbital and intrinsic spins always will commute with each other, $\left[\widehat{L}_{i}, \widehat{S}_{j}\right]=0$, since they are independent operators acting in different Hilbert spaces. Futhermore, all components of the total spin $\widehat{\mathbf{J}}$ commute with the spin-orbit operator $\widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}$. This follows from

$$
\left[\widehat{L}_{i}, \widehat{L}_{j} \widehat{S}_{j}\right]=\left[\widehat{L}_{i}, \widehat{L}_{j}\right] \widehat{S}_{j}=i \hbar \epsilon_{i j k} \widehat{L}_{k} \widehat{S}_{j}
$$

and

$$
\left[\widehat{S}_{i}, \widehat{L}_{j} \widehat{S}_{j}\right]=\widehat{L}_{j}\left[\widehat{S}_{i}, \widehat{S}_{j}\right]=i \hbar \epsilon_{i j k} \widehat{L}_{j} \widehat{S}_{k}
$$

Adding these two results and using the anti-symmetry of the Levi-Civita symbol, the two terms on the right-hand side cancel and thus $\left[\widehat{J}_{i}, \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}\right]=0$. This implies that also $\left[\widehat{\mathbf{J}}^{2}, \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}\right]=0$ in addition to $\left[\widehat{\mathbf{L}}^{2}, \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}\right]=\left[\widehat{\mathbf{S}}^{2}, \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}\right]=0$. Since both $\widehat{J}_{z}$ and the squared total spin

$$
\begin{equation*}
\widehat{\mathbf{J}}^{2}=\widehat{\mathbf{L}}^{2}+\widehat{\mathbf{S}}^{2}+2 \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}} \tag{8.17}
\end{equation*}
$$

thus commute with the spin-orbit opertor $\widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}$, we know that the eigenstates of these spin operators also will be eigenstates of the spin-orbit Hamiltonian (8.12). It will therefore be diagonal in the basis provided by these eigenstates defined by

$$
\begin{align*}
\widehat{\mathbf{J}}^{2}|j, m\rangle & =\hbar^{2} j(j+1)|j, m\rangle  \tag{8.18}\\
\widehat{J}_{z}|j, m\rangle & =\hbar m|j, m\rangle \tag{8.19}
\end{align*}
$$

Since $\widehat{J}_{z}=\widehat{L}_{z}+\widehat{S}_{z}$, the three azimuthal quantum numbers will always be related by $m=m_{l}+m_{s}$ where $\widehat{L}_{z}$ and $\widehat{S}_{z}$ act on the previous product states $\left|m_{l} ; m_{s}\right\rangle$.
In this new basis it is now straightforward to find the eigenvalues of the spin-orbit operator $2 \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}$. From (8.17) we have $2 \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}=\widehat{\mathbf{J}}^{2}-\widehat{\mathbf{L}}^{2}-\widehat{\mathbf{S}}^{2}$. Since the basis states $|j, m\rangle$ are also eigenstates of $\widehat{\mathbf{L}}^{2}$ and $\widehat{\mathbf{S}}^{2}$, we get

$$
\begin{equation*}
2 \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}|j, m\rangle=\hbar^{2}[j(j+1)-l(l+1)-1 / 2(1 / 2+1)]|j, m\rangle \tag{8.20}
\end{equation*}
$$

With $j=l \pm 1 / 2$ we therefore immediately have the eigenvalues

$$
2 \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}=\hbar^{2} \begin{cases}+l, & j=l+1 / 2  \tag{8.21}\\ -(l+1), & j=l-1 / 2\end{cases}
$$

in this new basis. They are obviously the same as those we found in the previous section by diagonalization of the $2 \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}$ operator in the $\left|m_{l} ; m_{s}\right\rangle$ basis.

These states are now straightforward to find for the case under consideration with intrin$\operatorname{sic} \operatorname{spin} s=1 / 2$. Again we start with the highest state $|l,+1 / 2\rangle$. Since here $m=l+1 / 2$, this state has also total angular momentum $j=l+1 / 2$. In the $|j, m\rangle$ basis we therefore have

$$
\begin{equation*}
|l+1 / 2, l+1 / 2\rangle=|l ;+1 / 2\rangle \tag{8.22}
\end{equation*}
$$

The next state in this ladder is then obtained by applying the lowering operator $\widehat{J}_{-}$on the left of (8.22) and $\widehat{L}_{-}+\widehat{S}_{-}$on the right-hand side. The results of these different operations all follow from the master formula (5.70). It gives

$$
\sqrt{2 l+1}|l+1 / 2, l-1 / 2\rangle=\sqrt{2 l}|l-1 ;+1 / 2\rangle+|l ;-1 / 2\rangle
$$

and therefore for the normalized state

$$
\begin{equation*}
|l+1 / 2, l-1 / 2\rangle=\sqrt{\frac{2 l}{2 l+1}}|l-1 ;+1 / 2\rangle+\sqrt{\frac{1}{2 l+1}}|l ;-1 / 2\rangle \tag{8.23}
\end{equation*}
$$

Applying $\widehat{J}_{-}=\widehat{L}_{-}+\widehat{S}_{-}$once more gives the next state down the spin ladder. It is

$$
\begin{equation*}
|l+1 / 2, l-3 / 2\rangle=\sqrt{\frac{2 l-1}{2 l+1}}|l-2 ;+1 / 2\rangle+\sqrt{\frac{2}{2 l+1}}|l-1 ;-1 / 2\rangle \tag{8.24}
\end{equation*}
$$

Continuing this way, we can reach all the states down to the lowest. It is just the reflection of the highest state (8.22), i.e. $|l+1 / 2,-l-1 / 2\rangle=|-l ;-1 / 2\rangle$. Any state in this ladder can be written as

$$
\begin{align*}
|l+1 / 2, m\rangle & =\sqrt{\frac{l+m+1 / 2}{2 l+1}}|m-1 / 2 ;+1 / 2\rangle \\
& +\sqrt{\frac{l-m+1 / 2}{2 l+1}}|m+1 / 2 ;-1 / 2\rangle \tag{8.25}
\end{align*}
$$

They all have total angular momentum $j=l+1 / 2$ and there are $2 j+1=2 l+2$ of them.

In order to construct the remaining $2 l$ states, we observe that there is just one other state orthogonal to (8.23). It is

$$
\begin{equation*}
|l-1 / 2, l-1 / 2\rangle=-\sqrt{\frac{1}{2 l+1}}|l-1 ;+1 / 2\rangle+\sqrt{\frac{2 l}{2 l+1}}|l ;-1 / 2\rangle \tag{8.26}
\end{equation*}
$$

and has total azimuthal quantum number $m=l-1 / 2$. Since there are no other states now with a higher value, this must be the highest state in a new spin ladder with total spin $j=l-1 / 2$. The remaining states in this ladder can again be obtained by applying $\widehat{J}_{-}$. But we know that the results are all orthogonal to the general $j=l+1 / 2$ state (8.25) and are therefore given by

$$
\begin{align*}
|l-1 / 2, m\rangle & =-\sqrt{\frac{l-m+1 / 2}{2 l+1}}|m-1 / 2 ;+1 / 2\rangle \\
& +\sqrt{\frac{l+m+1 / 2}{2 l+1}}|m+1 / 2 ;-1 / 2\rangle \tag{8.27}
\end{align*}
$$

It should by now be clear that the eigenstates (8.25) and (8.27) are just the same as the previously found states in (8.15) and (8.16) since $m=m_{l}+m_{s}$. Their components in the $\left|m_{l} ; m_{s}\right\rangle$ basis are called Clebsch-Gordan coefficients. For the addition of spin $l$ to spin $s=1 / 2$, these can now be read off from (8.25) and (8.27).

Addition of spins with different values and construction of the corresponding eigenstates, can be performed exactly along the same lines as above. For example adding intrinsic spin $s=1$ to orbital spin $l \geq 1$ will give $3(2 l+1)=6 l+3$ states in total. The highest state has $m=l+1$ and therefore total spin $j=l+1$. In this spin ladder there will thus be $2 j+1=2 l+3$ states. After applying $\widehat{J}_{-}$once, we find the state $|l+1, l\rangle$. As above, there will be another state orthogonal to it, i.e. the state $|l, l\rangle$. This will be the highest state in a new ladder with total angular momentum $l$ and therefore $2 l+1$ states altogether. We thus have constructed $2 l+3+2 l+1=4 l+4$ new eigenstates. But this number falls short of the total number $6 l+3$ states we started out with. The reason is now that when we apply $\widehat{J}_{-}$to the state $|l,+l\rangle$, we generate the eigenstate $|l, l-1\rangle$. But together with $|l+1, l-1\rangle$ these two states don't make up all the states with $m=l-1$. There is now a third one $|l-1, l-1\rangle$, orthogonal to these two. It will be the highest state in a new ladder with $j=l-1$ and therefore containing $2 l-1$ states. In this way we have found all the $6 l+3$ eigenstates. In summary, adding spin $l$ to spin $s=1$ gives total spins $j=(l+1, l, l-1)$.

This result is easily generalized to the addition of arbitrary spins $l$ and $s \leq l$. There will result a total spin with values $j=(l+s, l+s-1, \ldots, l-s)$. In a classical picture, the maximum value $l+s$ corresponds to the two spin vectors pointing in the same direction while the minimum value $l-s$ happens when they are in opposite directions. There are now in total $(2 s+1)(2 l+1)$ states which split up into these new spin multiplets, each with $2 j+1$ sub-states. The total number of these new spin eigenstates is therefore

$$
\begin{align*}
\sum_{j=l-s}^{l+s}(2 j+1) & =\sum_{k=0}^{2 s}[2(l-s+k)+1]=[2(l-s)+1](2 s+1)+2 s(2 s+1) \\
& =(2 s+1)(2 l+1) \tag{8.28}
\end{align*}
$$

as it should be. For each value of $l$ and $s$ one can calculate the corresponding ClebschGordan coefficients as we have already done for $s=1 / 2$. They can also be looked up in many textbooks and physical tables.

As a concrete example, lets work out the addition of $s=1 / 2$ and $l=1$. This comprises $2 \times 3=6$ product states of which the highest is $|3 / 2,+3 / 2\rangle=|+1 ;+1 / 2\rangle$. We have here already indicated that it has total spin $j=3 / 2$. We then apply the lowering operator $\widehat{J}_{-}=\widehat{L}_{-}+\widehat{S}_{-}$on both sides with the result $\sqrt{3}|3 / 2,+1 / 2\rangle=|+1 ;-1 / 2\rangle+$ $\sqrt{2}|0 ;+1 / 2\rangle$ from which we can read off the new eigenstate $|3 / 2,+1 / 2\rangle$. Applying now $\widehat{J}_{-}$to this state, gives similarly $2|3 / 2,-1 / 2\rangle=\sqrt{8 / 3}|0 ;-1 / 2\rangle+\sqrt{4 / 3}|-1 ;+1 / 2\rangle$ from which $|3 / 2,-1 / 2\rangle$ follows. The last state in this ladder can be obtained by using $\hat{J}_{-}$ once more with the result $|3 / 2,-3 / 2\rangle=|-1 ;-1 / 2\rangle$ which we also could have written down without any calculation. We thus find the quartet of spin- $3 / 2$ states:

$$
\begin{aligned}
|3 / 2,+3 / 2\rangle & =|+1 ;+1 / 2\rangle \\
|3 / 2,+1 / 2\rangle & =\sqrt{\frac{1}{3}}|+1 ;-1 / 2\rangle+\sqrt{\frac{2}{3}}|0 ;+1 / 2\rangle
\end{aligned}
$$

$$
\begin{align*}
|3 / 2,-1 / 2\rangle & =\sqrt{\frac{2}{3}}|0 ;-1 / 2\rangle+\sqrt{\frac{1}{3}}|-1 ;+1 / 2\rangle \\
|3 / 2,-3 / 2\rangle & =|-1 ;-1 / 2\rangle \tag{8.29}
\end{align*}
$$

The two states in the $j=1 / 2$ doublet now follows from orthogonality with the two quartet states with $z$-components $m= \pm 1 / 2$ :

$$
\begin{align*}
|1 / 2,+1 / 2\rangle & =\sqrt{\frac{2}{3}}|+1 ;-1 / 2\rangle-\sqrt{\frac{1}{3}}|0 ;-1 / 2\rangle  \tag{8.30}\\
|1 / 2,-1 / 2\rangle & =\sqrt{\frac{1}{3}}|0 ;-1 / 2\rangle-\sqrt{\frac{2}{3}}|-1 ;+1 / 2\rangle \tag{8.31}
\end{align*}
$$

Althogether we then have a new basis involving $4+2=6$ eigenstates of total angular momentum and have thus exhausted all the 6 available product states. As a bonus, we can now read off directly the explicit Clebsch-Gordan coefficients for adding spin- $1 / 2$ to spin-1.

### 8.4 Fine structure

Let us now return to the spin-orbit Hamiltonian (8.11) for the electron in the hydrogen atom. It is no longer described by the wave functions (5.141) in the $\left|n l m_{l} ; m_{s}\right\rangle$ basis. Instead we now have new, spinorial wave functions

$$
\begin{equation*}
\Psi_{n l j m}(r, \theta, \phi)=R_{n l}(r) \mathcal{Y}_{j m}(\theta, \phi) \tag{8.32}
\end{equation*}
$$

with the same radial functions $R_{n l}$. The previous eigenstates $Y_{l m}$ of orbital angular momentum $\widehat{\mathbf{L}}$ are replaced by the spinor eigenfunctions $\mathcal{Y}_{j m}$ of total angular momentum $\widehat{\mathbf{J}}=\widehat{\mathbf{L}}+\widehat{\mathbf{S}}$. For $j=l+1 / 2$ these can be read off (8.25) and become

$$
\begin{equation*}
\mathcal{Y}_{j m}=\sqrt{\frac{l+m+1 / 2}{2 l+1}} Y_{l, m-1 / 2} \alpha+\sqrt{\frac{l-m+1 / 2}{2 l+1}} Y_{l, m+1 / 2} \beta \tag{8.33}
\end{equation*}
$$

where $\alpha$ and $\beta$ are the eigenspinors for intrinsic spin up or down as introduced in Chapter 5. Similarly, for $j=l-1 / 2$ from (8.27) follows

$$
\begin{equation*}
\mathcal{Y}_{j m}=-\sqrt{\frac{l-m+1 / 2}{2 l+1}} Y_{l, m-1 / 2} \alpha+\sqrt{\frac{l+m+1 / 2}{2 l+1}} Y_{l, m+1 / 2} \beta \tag{8.34}
\end{equation*}
$$

Needless to say, they are eigenstates of the spin-orbit operator $2 \widehat{\mathbf{L}} \cdot \widehat{\mathbf{S}}$ with the eigenvalues (8.21). The energy level shift caused by this coupling (8.11) follows now directly from

$$
\begin{equation*}
E_{L S}=\int d^{3} x \Psi_{n l j m}^{\dagger} H_{L S} \Psi_{n l j m}=\frac{Z \alpha \hbar}{4 m_{e}^{2} c}\left\langle\frac{1}{r^{3}}\right\rangle\langle 2 \mathbf{L} \cdot \mathbf{S}\rangle \tag{8.35}
\end{equation*}
$$

The radial expectation value here is again given by the integral (8.4) which is

$$
\begin{equation*}
\left\langle\frac{a^{3}}{r^{3}}\right\rangle=\frac{2}{n^{3} l(l+1)(2 l+1)} \tag{8.36}
\end{equation*}
$$

for the power $s=3$. Here $a=a_{0} / Z$ is the ground-state radius of the atom in terms of the Bohr radius $a_{0}=\hbar / \alpha m_{e} c$. The expectation value of the spin-orbit operator follows now directly from (8.20) with the final result

$$
\begin{equation*}
E_{L S}=\frac{(Z \alpha)^{4}}{2 n^{3}} m_{e} c^{2}\left[\frac{j(j+1)-l(l+1)-3 / 4}{l(l+1)(2 l+1)}\right] \tag{8.37}
\end{equation*}
$$

The states with $j=l+1 / 2$ will be shifted up and the states with $j=l-1 / 2$ are pushed slightly more down with respect to the unperturbed levels as sketched in Fig.8.1. Comparing with the relativistic velocity shift (8.8), we see that these two shifts have similar absolute magnitudes. That should not be so surprising since the spin-orbit coupling arising from the Biot-Savart law is also a first-order effect in relativistic electrodynamics.

Needless to say, the spin-orbit shift formula (8.37) is not valid for $l=0$, i.e. $s$-states. For these there are no shifts since there is no orbital angular momentum to couple to. However, there is an extra contribution to these states called the Darwin term. It is a subtle quantum effect not easily explained in a quantitative way but comes naturally out of the Dirac equation for the electron. It will perhaps derived in a later chapter when we consider relativistic quantum mechanics.

Combining now the velocity and spin-orbit perturbations, we get for fine structure shift $E_{f i n e}=E_{v e l}+E_{L S}$ the simpler result

$$
\begin{align*}
E_{\text {fine }} & =\frac{(Z \alpha)^{4}}{2 n^{3}} m_{e} c^{2}\left[\frac{j(j+1)-l(l+1)-3 / 4}{l(l+1)(2 l+1)}-\frac{2}{2 l+1}+\frac{3}{4 n}\right] \\
& =-\left(\frac{Z \alpha}{n}\right)^{4} m_{e} c^{2}\left[\frac{n}{2 j+1}-\frac{3}{8}\right] \tag{8.38}
\end{align*}
$$

for a general energy level specified by the quantum numbers $n l_{j}$. It is valid both for $j=l+1 / 2$ and for $j=l-1 / 2$. It is now also valid for the $s$-states including the Darwin term. We notice that it has exactly the same form as the velocity perturbation (8.8) but with the replacement $l \rightarrow j$. This indicates that there must be a deeper theory underneath our treatment here. An unperturbed level whose energy is specified by the principal quantum number $n$ alone, will therefore split up into several components. Those with same value of total spin $j$, remain degenerate. For example, the $n=2$ level gives rise to the perturbed states $2 s_{1 / 2}, 2 p_{1 / 2}$ and $2 p_{3 / 2}$ which are all shifted as seen in Fig.8.2. But $2 s_{1 / 2}$ and $2 p_{1 / 2}$ are shifted by the same amount and therefore stay degenerate. Similarly, at $n=3$ we get the perturbed states $3 s_{1 / 2}, 3 p_{1 / 2}, 3 p_{3 / 2}, 3 d_{3 / 2}$ and $3 d_{5 / 2}$. Here $3 s_{1 / 2}$ and $3 p_{1 / 2}$ remain degenerate as well as $3 p_{3 / 2}$ and $3 d_{3 / 2}$. This level is thus split into three sublevels where $3 d_{5 / 2}$ has the highest energy as follows from (8.38).
It should be obvious that the fine structure also will give rise to a more complex optical spectrum from radiative transitions between these levels. For instance, the leading line $H_{\alpha}$ in the Balmer series arises from the $(n=3) \rightarrow(n=2)$ transition and will now be split up into several nearby lines determined by the selection rules. It was accurate measurements of the frequencies of all these lines that historically showed that quantum mechanics was the correct theory. However, with better experimental methods it was found that the degeneracy between the $2 s_{1 / 2}$ and $2 p_{1 / 2}$ levels was not perfect, but corresponded to a frequency of 1057 MHz . This is around $10 \%$ of the splitting between


Figure 8.2: Fine structure in the hydrogen spectrum. (not to scale)
the $j=1 / 2$ and $j=3 / 2$ levels and is called the Lamb shift. It can only be understood within QED which also explained the anomalous $g$-factor for the electron.

### 8.5 Anomalous Zeeman effect and Landé g-factor

Let us now consider again the Zeeman effect when the electron is in an external magnetic field $\mathbf{B}$. In the previous chapter it was calculated without caring about the fine structure of the atomic states. That result therefore applies only in very strong fields $B \gg 1 \mathrm{~T}$ which overwhelms the relativistic effects. When the field is much weaker, the atom is described by the new eigenstates $n l_{j}$ of total angular momentum. When the field is along the $z$-axis, the interaction Hamiltonian (7.35) gives the expectation value

$$
\begin{equation*}
\left\langle\widehat{H}_{B}\right\rangle=E_{B}=\frac{e}{2 m_{e}}\langle j, m| \widehat{L}_{z}+2 \widehat{S}_{z}|j, m\rangle B \tag{8.39}
\end{equation*}
$$

for the energy. The magnetic moment operator (7.31) is not diagonal in this basis. However, since $\widehat{J}_{z}=\widehat{L}_{z}+\widehat{S}_{z}$ and $\widehat{J}_{z}|j, m\rangle=\hbar m|j, m\rangle$, it simplifies to

$$
\begin{equation*}
E_{B}=\frac{e}{2 m_{e}}\left(m \hbar+\langle j, m| \widehat{S}_{z}|j, m\rangle\right) B \tag{8.40}
\end{equation*}
$$

This last term gives rise to what is called the anomalous Zeeman effect and is caused by the intrinsic spin.

We obtain the needed matrix element from the states $|j, m\rangle$ derived in (8.25) and (8.27). In the case $j=l+1 / 2$ we find

$$
\begin{aligned}
\widehat{S}_{z}|j, m\rangle & =(+1 / 2 \hbar) \sqrt{\frac{l+m+1 / 2}{2 l+1}}|m-1 / 2 ;+1 / 2\rangle \\
& +(-1 / 2 \hbar) \sqrt{\frac{l-m+1 / 2}{2 l+1}}|m+1 / 2 ;-1 / 2\rangle
\end{aligned}
$$

and therefore

$$
\begin{equation*}
\langle j, m| \widehat{S}_{z}|j, m\rangle=\hbar \frac{(l+m+1 / 2)(+1 / 2)+(l-m+1 / 2)(-1 / 2)}{2 l+1}=\frac{\hbar m}{2 l+1} \tag{8.41}
\end{equation*}
$$

Similarly, when $j=l-1 / 2$ the same result is obtained, but with opposite sign. As a result, we can write $E_{B}=g \mu_{B} B m$ where

$$
g=1 \pm \frac{1}{2 l+1}=\frac{2}{2 l+1} \times \begin{cases}l+1, & j=l+1 / 2  \tag{8.42}\\ l, & j=l-1 / 2\end{cases}
$$

is called the Landé $g$-factor. We can write the matrix element (8.41) more compactly as $\langle j, m| \widehat{S}_{z}|j, m\rangle=2 \hbar m(j-l) /(2 l+1)$ so that the result for the $g$-factor becomes

$$
\begin{equation*}
g=\frac{2 j+1}{2 l+1} \tag{8.43}
\end{equation*}
$$

It is $g=2$ for the $2 s_{1 / 2}$ level as for a free electron, $g=2 / 3$ for $2 p_{1 / 2}$ and $g=4 / 3$ for the $2 p_{3 / 2}$ level.
It is customary to define the measured magnetic moment $\mu$ of the atom as given by the above interaction energy when the azimuthal quantum number has the maximum value $m=j$. In a semi-classical picture this corresponds to having the total spin vector along the $z$-axis. Then we can write $E_{B}=\mu B$ with the magnetic moment $\mu=g j \mu_{B}$. One can also define an effective magnetic moment vector operator

$$
\begin{equation*}
\widehat{\boldsymbol{\mu}}_{e f f}=\frac{e g}{2 m_{e}} \widehat{\mathbf{J}} \tag{8.44}
\end{equation*}
$$

which has the same expectation values as found above. Semi-classically, it points in the same direction as the total angular momentum.
The result (8.41) is an example of a more general result which is summarized in the Wigner-Eckart theorem. It can be proven for any vector operator $\widehat{V}_{i}$ which is defined to have the commutator

$$
\begin{equation*}
\left[\widehat{J}_{i}, \widehat{V}_{j}\right]=i \hbar \epsilon_{i j k} \widehat{V}_{k} \tag{8.45}
\end{equation*}
$$

From the beginning of Chapter 5 we know that both the position operator $\widehat{x}_{i}$ and momentum operator $\widehat{p}_{i}$ are such operators. Then it follows that also the orbital spin $\widehat{L}_{i}$ is such a vector operator as is the intrinsic spin $\widehat{S}_{i}$. For all such vector operators the theorem says that the matrix element $\langle j, m| \widehat{\mathbf{V}}|j, m\rangle$ is proportional to the matrix element $\langle j, m| \widehat{\mathbf{J}}|j, m\rangle$. More formally, this can be summarized in the equation

$$
\begin{equation*}
\langle j, m| \widehat{\mathbf{V}}|j, m\rangle=V_{j j}\langle j, m| \widehat{\mathbf{J}}|j, m\rangle \tag{8.46}
\end{equation*}
$$

where $V_{j j}$ is called the reduced matrix element. It is independent of the azimuthal quantum number $m$. We have verified it for the operator $\widehat{S}_{z}$. A similar calculation will also show that it holds for matrix elements of $\widehat{L}_{z}$.

Accepting the theorem, we then have $\langle j, m| \widehat{S}_{z}|j, m\rangle=\hbar m S_{j j}$ where the reduced spin matrix element now can be obtained from $\langle j, j| \widehat{S}_{z}|j, j\rangle=\hbar j S_{j j}$. This is achieved by
writing $2 \widehat{\mathbf{J}} \cdot \widehat{\mathbf{S}}=2 \widehat{J}_{z} \widehat{S}_{z}+\widehat{J}_{-} \widehat{S}_{+}+\widehat{J}_{+} \widehat{S}_{-}=2 \widehat{J}_{z} \widehat{S}_{z}+\widehat{J}_{-} \widehat{S}_{+}+\widehat{S}_{-} \widehat{J}_{+}+2 \hbar \widehat{S}_{z}$ where we in the last equality make use of the commutator $\left[\widehat{J}_{+}, \widehat{S}_{-}\right]=\left[\widehat{S}_{+}, \widehat{S}_{-}\right]=2 \hbar \widehat{S}_{z}$. Thus

$$
\langle j, j| \widehat{\mathbf{J}} \cdot \widehat{\mathbf{S}}|j, j\rangle=\hbar(j+1)\langle j, j| \widehat{S}_{z}|j, j\rangle
$$

since the terms involving $\widehat{J}_{-}$and $\widehat{J}_{+}$give zero when sandwiched between highest states. On the other hand, we have $2 \widehat{\mathbf{J}} \cdot \widehat{\mathbf{S}}=\widehat{\mathbf{J}}^{2}+\widehat{\mathbf{S}}^{2}-\widehat{\mathbf{L}}^{2}$ from squaring $\widehat{\mathbf{L}}=\widehat{\mathbf{J}}-\widehat{\mathbf{S}}$. Combining these two results, it follows that

$$
\begin{equation*}
S_{j j}=\frac{j(j+1)+s(s+1)-l(l+1)}{2 j(j+1)} \tag{8.47}
\end{equation*}
$$

This now also gives the matrix element $\langle j, m| \widehat{S}_{z}|j, m\rangle$. From (8.39) we can then read off original expression for the Landé $g$-factor,

$$
\begin{equation*}
g=1+\frac{j(j+1)+s(s+1)-l(l+1)}{2 j(j+1)} \tag{8.48}
\end{equation*}
$$

Taking here $j=l \pm 1 / 2$, we recover the previous version (8.42) of the same result.
With the result (8.47) for the reduced spin matrix element $S_{j j}$ we can write the WignerEckart theorem (8.46) for this operator as

$$
\begin{equation*}
\langle j, m| \widehat{\mathbf{S}}|j, m\rangle=\frac{\langle j, m| \widehat{\mathbf{S}} \cdot \widehat{\mathbf{J}}|j, m\rangle}{j(j+1) \hbar^{2}}\langle j, m| \widehat{\mathbf{J}}|j, m\rangle \tag{8.49}
\end{equation*}
$$

On this form the theorem is in fact valid for any vector operator $\widehat{\mathbf{V}}$. To prove it, consider the matrix element $\langle j, m| \widehat{\mathbf{V}} \cdot \widehat{\mathbf{J}}|j, m\rangle$ and insert a complete set of total angular momentum states between the two operators. Then

$$
\begin{aligned}
\langle j, m| \widehat{\mathbf{V}} \cdot \widehat{\mathbf{J}}|j, m\rangle & =\sum_{m^{\prime}=-j}^{j}\langle j, m| \widehat{\mathbf{V}}\left|j, m^{\prime}\right\rangle \cdot\left\langle j, m^{\prime}\right| \widehat{\mathbf{J}}|j, m\rangle \\
& =V_{j j} \sum_{m^{\prime}=-j}^{j}\langle j, m| \widehat{\mathbf{J}}\left|j, m^{\prime}\right\rangle \cdot\left\langle j, m^{\prime}\right| \widehat{\mathbf{J}}|j, m\rangle
\end{aligned}
$$

using the definition (8.46) of the theorem. Now we undo the completeness sum and obtain

$$
\langle j, m| \widehat{\mathbf{V}} \cdot \widehat{\mathbf{J}}|j, m\rangle=V_{j j}\langle j, m| \widehat{\mathbf{J}} \cdot \widehat{\mathbf{J}}|j, m\rangle=V_{j j} j(j+1) \hbar^{2}
$$

Thus we have the reduced matrix element $V_{j j}$ and therefore Wigner-Eckart theorem on the very useful form

$$
\begin{equation*}
\langle j, m| \widehat{\mathbf{V}}|j, m\rangle=\frac{\langle j, m| \widehat{\mathbf{V}} \cdot \widehat{\mathbf{J}}|j, m\rangle}{j(j+1) \hbar^{2}}\langle j, m| \widehat{\mathbf{J}}|j, m\rangle \tag{8.50}
\end{equation*}
$$

It will serve us well in the next section.

### 8.6 Hyperfine splitting

The spin-orbit interaction can be understood by the motion of the nuclear electric charge creating a magnetic field which couples to the spin of the electron. Considering a nucleus with charge $Z e$ and mass $m_{N}$, we write the magnetic moment as

$$
\begin{equation*}
\boldsymbol{\mu}_{N}=\frac{Z e g_{N}}{2 m_{N}} \mathbf{I} \tag{8.51}
\end{equation*}
$$

where $g_{N}$ is the nuclear $g$-factor and $\mathbf{I}$ the spin of the nucleus. For a single proton we had $g_{p} / 2=2.78$. The magnetic moment creates a vector potential

$$
\begin{equation*}
\mathbf{A}=\frac{\mu_{0}}{4 \pi} \boldsymbol{\mu}_{N} \wedge \frac{\mathbf{r}}{r^{3}}=-\frac{1}{4 \pi \epsilon_{0} c^{2}} \boldsymbol{\mu}_{N} \wedge \nabla \frac{1}{r} \tag{8.52}
\end{equation*}
$$

and thus a magnetic field $\mathbf{B}=\boldsymbol{\nabla} \wedge \mathbf{B}$. It gives rise to a standard magnetic coupling $H_{m a g}=-\boldsymbol{\mu}_{e} \cdot \mathbf{B}$ to the electron. Making use of the vector identity

$$
\boldsymbol{\nabla} \wedge(\mathbf{V} \wedge \nabla S)=\mathbf{V} \nabla^{2} S-\boldsymbol{\nabla}(\mathbf{V} \cdot \boldsymbol{\nabla} S)
$$

it is found to be

$$
H_{\text {mag }}=\frac{1}{4 \pi \epsilon_{0} c^{2}}\left[\left(\boldsymbol{\mu}_{e} \cdot \boldsymbol{\mu}_{N}\right) \nabla^{2} \frac{1}{r}-\left(\boldsymbol{\mu}_{e} \cdot \boldsymbol{\nabla}\right)\left(\boldsymbol{\mu}_{N} \cdot \boldsymbol{\nabla}\right) \frac{1}{r}\right]
$$

The last term here can now be simplified using

$$
\partial_{i} \partial_{j} \frac{1}{r}=3 \frac{x_{i} x_{j}}{r^{5}}-\delta_{i j}\left(\frac{1}{r^{3}}-\frac{1}{3} \nabla^{2} \frac{1}{r}\right)
$$

Collecting terms, we thus have

$$
\begin{equation*}
H_{m a g}=\frac{1}{4 \pi \epsilon_{0} c^{2}}\left(\frac{2}{3} \boldsymbol{\mu}_{e} \cdot \boldsymbol{\mu}_{N} \nabla^{2} \frac{1}{r}+\frac{1}{r^{3}}\left[\boldsymbol{\mu}_{e} \cdot \boldsymbol{\mu}_{N}-3 \frac{\left(\boldsymbol{\mu}_{e} \cdot \mathbf{r}\right)\left(\boldsymbol{\mu}_{N} \cdot \mathbf{r}\right)}{r^{2}}\right]\right) \tag{8.53}
\end{equation*}
$$

We recognize the last two terms as the dipole-dipole interaction between the proton and electron magnetic moments. It is also called a tensor interaction and will not contribute to states with orbital angular momentum $l>0$.

For $s$-states only the first term gives a non-zero result. We can write this part as

$$
\begin{equation*}
H_{F}=-\frac{1}{4 \pi \epsilon_{0} c^{2}} \frac{8 \pi}{3} \boldsymbol{\mu}_{e} \cdot \boldsymbol{\mu}_{N} \delta(\mathbf{r}) \tag{8.54}
\end{equation*}
$$

when we make use of the relation

$$
\begin{equation*}
\nabla^{2} \frac{1}{r}=-4 \pi \delta(\mathbf{r}) \tag{8.55}
\end{equation*}
$$

It is just the Poisson equation for a unit point charge satisfying Coulomb's law. Since the magnetic moment of the electron is $\boldsymbol{\mu}_{e}=-\left(e g_{e} / 2 m_{e}\right) \mathbf{S}$ with electric charge which is now $-e$, we thus get

$$
\begin{equation*}
H_{F}=\frac{Z e^{2}}{4 \pi \epsilon_{0} c^{2}} \frac{g_{N}}{2 m_{e} m_{N}} \frac{8 \pi}{3} \delta(\mathbf{r}) \mathbf{I} \cdot \mathbf{S} \tag{8.56}
\end{equation*}
$$

when we set $g_{e}=2$ for the electron. This is the Fermi Hamiltonian describing the hyperfine interaction of $l=0$ states in the hydrogen atom.

Let us now calculate the expectation value of the Fermi operator (8.56). The spin part $\widehat{H}=2 \widehat{\mathbf{I}} \cdot \widehat{\mathbf{S}}$ is of exactly the same type as for the spin-orbit interaction previously considered in this chapter. It is simplest evaluated in a basis where this operator is diagonal, i.e. formed by the eigenstates of the total spin operator $\widehat{\mathbf{F}}=\widehat{\mathbf{I}}+\widehat{\mathbf{S}}$. Denoting these states by $\left|f, m_{f}\right\rangle$, we thus have

$$
\begin{align*}
\widehat{\mathbf{F}}^{2}\left|f, m_{f}\right\rangle & =\hbar^{2} f(f+1)\left|f, m_{f}\right\rangle  \tag{8.57}\\
\widehat{F}_{z}\left|f, m_{f}\right\rangle & =\hbar m_{f}\left|f, m_{f}\right\rangle \tag{8.58}
\end{align*}
$$

The highest of these new spin states is obviously $\left|m_{i}=+1 / 2\right\rangle\left|m_{s}=+1 / 2\right\rangle$ which we write as $|+1 / 2 ;+1 / 2\rangle \equiv|1,+1\rangle$ in the same notation as before when we added spins. We have already indicated that it corresponds to total spin $f=1$. Applying as before the lowering operator $\widehat{F}_{-}=\widehat{I}_{-}+\widehat{S}_{-}$on both sides, we construct the other states in the same ladder,

$$
\begin{align*}
|1,1\rangle & =|+1 / 2 ;+1 / 2\rangle \\
|1,0\rangle & =\sqrt{\frac{1}{2}}[|+1 / 2 ;-1 / 2\rangle+|-1 / 2 ;+1 / 2\rangle]  \tag{8.59}\\
|1,-1\rangle & =|-1 / 2 ;-1 / 2\rangle
\end{align*}
$$

These three states form a spin- 1 triplet. Of the original four product states there is one left. It must be orthogonal to $|1,0\rangle$ and is therefore

$$
\begin{equation*}
|0,0\rangle=\sqrt{\frac{1}{2}}[|+1 / 2 ;-1 / 2\rangle-|-1 / 2 ;+1 / 2\rangle] \tag{8.60}
\end{equation*}
$$

Obviously it must correspond to spin $f=0$ which also can be verified by a direct calculation. It is the spin-0 singlet state. In summary, adding spin $1 / 2$ to spin $1 / 2$ forms four states of which three constitute a spin-1 multiplet and one represents a spin0 state.

Instead of constructing these states in the abstract language of ket vectors, we could also use the spinor representation for spin $1 / 2$ introduced in Chapter 5 . The eigenstates for the electron spin $\mathbf{S}$ would then be the spinors $\alpha_{e}$ and $\beta_{e}$ while for the proton spin $\mathbf{I}$ it would be $\alpha_{p}$ and $\beta_{p}$. The triplet states (8.59) would then be written as

$$
\begin{align*}
T_{+} & =\alpha_{e} \alpha_{p} \\
T_{0} & =\sqrt{\frac{1}{2}}\left(\alpha_{e} \beta_{p}+\beta_{e} \alpha_{p}\right)  \tag{8.61}\\
T_{-} & =\beta_{e} \beta_{p}
\end{align*}
$$

while the singlet becomes

$$
\begin{equation*}
S=\sqrt{\frac{1}{2}}\left(\alpha_{e} \beta_{p}-\beta_{e} \alpha_{p}\right) \tag{8.62}
\end{equation*}
$$

Needless to say, these composite spin eigenstates for the total spin will be the same for the addition of any two spin- $1 / 2$ systems.

In this new basis the operator $2 \widehat{\mathbf{I}} \cdot \widehat{\mathbf{S}}=\widehat{\mathbf{F}}^{2}-\widehat{\mathbf{I}}^{2}-\widehat{\mathbf{S}}^{2}$ is now diagonal. Since both the nuclear $\operatorname{spin}$ is $i=1 / 2$ and the electron spin is $s=1 / 2$, it takes the values $\hbar^{2}[f(f+1)-3 / 4-3 / 4]$ or

$$
2 \widehat{\mathbf{I}} \cdot \widehat{\mathbf{S}}=\hbar^{2} \begin{cases}+1 / 2, & f=1  \tag{8.63}\\ -3 / 2, & f=0\end{cases}
$$

Each unperturbed energy level will therefore be split into two sublevels by an amount

$$
\begin{align*}
E_{F} & =\frac{Z e^{2}}{4 \pi \epsilon_{0} c^{2}} \frac{g_{N}}{2 m_{e} m_{N}} \frac{8 \pi}{3} \int d^{3} x \psi_{n l m}^{*} \delta(\mathbf{r}) \psi_{n l m}\langle\mathbf{I} \cdot \mathbf{S}\rangle \\
& =\frac{Z e^{2}}{4 \pi \epsilon_{0} c^{2}} \frac{g_{N}}{2 m_{e} m_{N}} \frac{8 \pi}{3}\left|\psi_{n 00}(0)\right|^{2}\langle\mathbf{I} \cdot \mathbf{S}\rangle \tag{8.64}
\end{align*}
$$

corresponding to (8.35) for the spin-orbit coupling. This is illustrated in Fig.8.3. Thus


Figure 8.3: Hyperfine splitting of a $s$-level in arbitrary units.
we see explicitly that only $l=0$ states will be affected by this hyperfine perturbation since states with $l>0$ vanish at the origin. For the $s$-states we make use of the result (5.149) which gives $\left|\psi_{n 00}(0)\right|^{2}=1 / \pi n^{3} a^{3}$. The atomic $1 s$ radius is again given in terms of the Bohr radius as $a=a_{0} / Z$. One thus obtains

$$
\begin{equation*}
E_{F}=\frac{2 g_{N}}{3 n^{3}}(Z \alpha)^{4} m_{e} c^{2}\left(\frac{m_{e}}{m_{N}}\right)[f(f+1)-3 / 2] \tag{8.65}
\end{equation*}
$$

Comparing with the spin-orbit result (8.37) we see that the hyperfine perturbation in hydrogen is much smaller by a factor $m_{e} / m_{p}=1 / 1836$. It is for this reason that it has earned its name. Notice that for this part of the hyperfine interaction we have the effective Hamiltonian

$$
\begin{equation*}
H_{F}^{\prime}=K_{0} \mathbf{I} \cdot \mathbf{S} \tag{8.66}
\end{equation*}
$$

where the coupling constant is seen to be given by

$$
\begin{equation*}
K_{0} \hbar^{2}=\frac{4 g_{N}}{3 n^{3}}(Z \alpha)^{4} m_{e} c^{2}\left(\frac{m_{e}}{m_{N}}\right) \tag{8.67}
\end{equation*}
$$

It is seen to decrease rapidly for higher excited states when the principal quantum number $n$ gets large.

The hyperfine split between the two sublevels arising from the lowest $1 s_{1 / 2}$ in the hydrogen atom is $\Delta E_{F}=E_{F}(f=1)-E_{F}(f=0)$ or

$$
\begin{equation*}
\Delta E_{F}=\left(4 g_{p} m_{e} / 3 m_{p}\right) \alpha^{4} m_{e} c^{2} . \tag{8.68}
\end{equation*}
$$

Putting in numbers, one finds $\Delta E_{F}=5.9 \times 10^{-6} \mathrm{eV}$. By a magnetic dipole transition as described in the previous chapter, a photon can be emitted when the atom is in the upper triplet state. The corresponding frequency is $\nu=\Delta E_{F} / h=1420 \mathrm{MHz}$ or wavelength $\lambda=c / \nu=21.1 \mathrm{~cm}$. This corresponds to radiation in the radio band. Its detection has been very useful in mapping out the presence of neutral hydrogen atoms in the Universe.

In this calculation we have not so far considered the ordinary spin-independent interaction $H_{e l}=\left(e / m_{e}\right) \mathbf{p} \cdot \mathbf{A}$. Both this electric coupling and the above magnetic coupling are present in the Pauli Hamiltonian (7.30) and will in general contribute to the hyperfine interaction. Together with the vector potential (8.52) it gives

$$
\begin{equation*}
H_{e l}=\frac{1}{4 \pi \epsilon_{0} c^{2} r^{3}} \frac{e}{m_{e}} \mathbf{p} \cdot \boldsymbol{\mu}_{N} \wedge \mathbf{r}=\frac{1}{4 \pi \epsilon_{0} c^{2} r^{3}} \frac{e}{m_{e}} \boldsymbol{\mu}_{N} \cdot \mathbf{L} \tag{8.69}
\end{equation*}
$$

where $\mathbf{L}=\mathbf{r} \wedge \mathbf{p}$ is the orbital angular momentum of the electron. It is seen to contribute only for states with $l>0$. We can therefore group it together with the tensor interaction in (8.53) to give the combined interaction Hamiltonian

$$
\begin{equation*}
H_{l>0}=\frac{Z e^{2}}{4 \pi \epsilon_{0} c^{2} r^{3}} \frac{g_{N}}{2 m_{e} m_{N}}\left[\mathbf{I} \cdot \mathbf{L}-\mathbf{I} \cdot \mathbf{S}+\frac{3}{r^{2}}(\mathbf{I} \cdot \mathbf{r})(\mathbf{S} \cdot \mathbf{r})\right] \tag{8.70}
\end{equation*}
$$

When calculating the perturbative effect on an energy level with $l \geq 1$, we can now use the Wigner-Eckardt theorem on the form (8.50) to simplify the result. For this purpose we introduce the vector operator

$$
\mathbf{B}^{\prime}=\left[\frac{1}{r^{3}}(\mathbf{L}-\mathbf{S})+\frac{3}{r^{5}} \mathbf{r}(\mathbf{S} \cdot \mathbf{r})\right]
$$

It allows us to write the Hamiltonian (8.70) simply as $H_{l>0}=K \mathbf{I} \cdot \mathbf{B}^{\prime}$ where the constant $K=\left(Z e^{2} / 4 \pi \epsilon_{0} c^{2}\right)\left(g_{N} / 2 m_{e} m_{N}\right)$. Then we have

$$
\langle j, m| \widehat{\mathbf{B}}^{\prime} \cdot \widehat{\mathbf{J}}|j, m\rangle=\langle j, m| \frac{1}{r^{3}}(\widehat{\mathbf{L}}-\widehat{\mathbf{S}}) \cdot \widehat{\mathbf{J}}+\frac{3}{r^{5}}(\mathbf{r} \cdot \widehat{\mathbf{J}})(\widehat{\mathbf{S}} \cdot \mathbf{r})|j, m\rangle
$$

But now $(\widehat{\mathbf{L}}-\widehat{\mathbf{S}}) \cdot \widehat{\mathbf{J}}=\widehat{\mathbf{L}}^{2}-\widehat{\mathbf{S}}^{2}$ since $\widehat{\mathbf{J}}=\widehat{\mathbf{L}}+\widehat{\mathbf{S}}$. But both $\widehat{\mathbf{L}}^{2}$ and $\widehat{\mathbf{S}}^{2}$ are diagonal in the $\left|f, m_{f}\right\rangle$ basis with eigenvalues $\hbar^{2} l(l+1)$ and $3 \hbar^{2} / 4$ respectively. Also $\mathbf{r} \cdot \widehat{\mathbf{J}}=\mathbf{r} \cdot \widehat{\mathbf{S}}$ since $\mathbf{r} \cdot \mathbf{L}=\mathbf{r} \cdot(\mathbf{r} \wedge \mathbf{p})=0$. Therefore one finds $(\mathbf{r} \cdot \widehat{\mathbf{J}})(\widehat{\mathbf{S}} \cdot \mathbf{r})=(\mathbf{r} \cdot \widehat{\mathbf{S}})^{2}$. Remembering now that the electron has spin $\mathbf{S}=(\hbar / 2) \boldsymbol{\sigma}$, we have simply that $(\mathbf{r} \cdot \widehat{\mathbf{J}})(\widehat{\mathbf{S}} \cdot \mathbf{r})=(\hbar / 4)^{2} r^{2}$. The above matrix element thus simplifies to

$$
\langle j, m| \widehat{\mathbf{B}}^{\prime} \cdot \widehat{\mathbf{J}}|j, m\rangle=\hbar^{2}\left\langle\frac{1}{r^{3}}\right\rangle\left[l(l+1)-\frac{3}{4}+\frac{3}{4}\right]
$$

With the help of the Wigner-Eckart theorem (8.50) we thus can also find the matrix element $\langle j, m| \widehat{\mathbf{B}}^{\prime}|j, m\rangle$. We then see that we can effectively use

$$
\begin{equation*}
\mathbf{B}^{\prime}=\frac{l(l+1)}{j(j+1)}\left\langle\frac{1}{r^{3}}\right\rangle \mathbf{J} \tag{8.71}
\end{equation*}
$$

in all matrix elements of this part of the hyperfine interaction. As expected, it is seen to not contribute when $l=0$. The orbital matrix element $\left\langle 1 / r^{3}\right\rangle$ is the same (8.36) as used in the calculation of the spin-orbit coupling. In this way we end up with the effective operator

$$
\begin{equation*}
H_{l>0}^{\prime}=K_{1} \mathbf{I} \cdot \mathbf{J} \tag{8.72}
\end{equation*}
$$

for this part of the hyperfine interaction. Its magnitude is determined by the coupling constant

$$
\begin{align*}
K_{1} \hbar^{2} & =\frac{Z e^{2}}{4 \pi \epsilon_{0} c^{2}} \frac{g_{N}}{2 m_{e} m_{N}} \frac{l(l+1)}{j(j+1)}\left\langle\frac{1}{r^{3}}\right\rangle \\
& =\frac{g_{N}}{n^{3}}(Z \alpha)^{4} m_{e} c^{2}\left(\frac{m_{e}}{m_{N}}\right) \frac{1}{j(j+1)(2 l+1)} \tag{8.73}
\end{align*}
$$

The spin part is diagonalized by constructing the eigenstates of the total angular momentum $\mathbf{F}=\mathbf{I}+\mathbf{J}$. When the nuclear spin $i=1 / 2$ as in hydrogen, all states with orbital angular momentum $l \geq 1$ will therefore also be split into two sublevels as was the case for the $l=0$ states. We can combine the two effective operators (8.66) and (8.72) into the complete operator

$$
\begin{equation*}
H_{H F}^{\prime}=K_{0} \mathbf{I} \cdot \mathbf{S}+K_{1} \mathbf{I} \cdot \mathbf{J} \tag{8.74}
\end{equation*}
$$

which sums up the effects of these hyperfine interactions. But it should be kept in mind that they always operate on different states.

### 8.7 Magnetic transitions in hyperfine doublet

It has already ben mentioned that the hyperfine splitting of the hydrogen ground state has important astrophysical implications since the photons from the triplet $\rightarrow$ singlet transitions can be detected with radio telescopes and thus can signal the presence of hydrogen in the Universe. But this depends on the lifetime for the transition to be large enough so that the atoms can remain in the excited triplet state over sufficiently long time scales.

This must be a magnetic transition since it involves the electron spin. It was calculated in the previous chapter and given by (7.111). We thus need to calculate the matrix elements $\boldsymbol{\sigma}_{f i}=\langle S| \boldsymbol{\sigma}|T\rangle$ between the triplet and singlet states in hydrogen given by (8.61) and (8.62). Writing now

$$
\begin{equation*}
\left.\left.\left.\left|\boldsymbol{\sigma}_{f i}\right|^{2}=2\left|\langle S| \sigma_{-}\right| T\right\rangle\left.\right|^{2}+2\left|\langle S| \sigma_{+}\right| T\right\rangle\left.\right|^{2}+\left|\langle S| \sigma_{z}\right| T\right\rangle\left.\right|^{2} \tag{8.75}
\end{equation*}
$$

where the spin raising and lowering matrices $\sigma_{ \pm}=\left(\sigma_{x} \pm i \sigma_{y}\right) / 2$ were derived already in (5.76). Acting on the fundamental spin- $1 / 2$ spinors (5.85), we see that the only nonzero contributions come from $\sigma_{-} \alpha=\beta$ and $\sigma_{+} \beta=\alpha$. In addition, we obviously have $\sigma_{z} \alpha=\alpha$ and $\sigma_{z} \beta=-\beta$.
Assuming now for instance that initially the atom is in the triplet state with $z$-component $m_{s}=+1$, we must calculate the matrix element $S^{\dagger} \sigma_{-} T_{+}=S^{\dagger} \sigma_{-} \alpha_{e} \alpha_{p}$. Since the spin
matrices only act on the electron spinors $\alpha_{e}$ and $\beta_{e}$, this equals $S^{\dagger} \beta_{e} \alpha_{p}$. With the singlet spinor from (8.62), we thus find $S^{\dagger} \sigma_{-} T_{+}=-1 / \sqrt{2}$. The two other terms in (8.75) gives nothing. More specifically, the second term is zero since $\sigma_{+} T_{+}=0$ while in the third term $S^{\dagger} T_{+}=0$ from orthogonality.

The full, squared matrix element is therefore in this case $\left|\boldsymbol{\sigma}_{f i}\right|^{2}=1$. We would have obtained the same if we had assumed the atom to be in one of the other possible initial states. For instance, if it was in the state with $m_{s}=-1$ instead, only the middle term in (8.75) would contribute with the same result. On the other hand, for the initial state $T_{0}$, only the last term is non-zero since $\sigma_{z} T_{0}=S$ and with $S^{\dagger} S=1$ therefore giving $\left|\boldsymbol{\sigma}_{f i}\right|^{2}=1$ again.
With this value for the matrix element, we now get the hyperfine transition rate

$$
\begin{equation*}
\Gamma_{h y p}=\frac{\alpha}{3} \omega\left(\frac{\hbar \omega}{m_{e} c^{2}}\right)^{2} \tag{8.76}
\end{equation*}
$$

The frequency $\omega$ follows from the energy split (8.68) and its numerical value gives the rate $\Gamma=2.85 \times 10^{-15} \mathrm{~s}^{-1}$ corresponding to a lifetime of $\tau=1 / \Gamma=3.50 \times 10^{14} \mathrm{~s}$. Since there are $3.16 \times 10^{7} \mathrm{~s}$ in one year, we see that the atom will remain in the triplet state for ten million years.

We can now also calculate the angulular distribution of the photons in this hyperfine transition along exactly the same lines as for the electric transition in the previous chapter. From (7.110) we have the differential transition rate

$$
\begin{equation*}
\frac{d \Gamma}{d \Omega}=\left(\frac{e \hbar}{2 m_{e}}\right)^{2} \frac{\omega}{8 \pi^{2} \epsilon_{0} \hbar c^{3}}\left|\mathbf{e}_{\lambda}^{*} \cdot\left(\mathbf{k} \wedge \boldsymbol{\sigma}_{f i}\right)\right|^{2} \tag{8.77}
\end{equation*}
$$

when the photon has the wave vector $\mathbf{k}$ and polarization $\mathbf{e}_{\lambda}$. If we quantize the spin in the initial state along the $z$-axis, $\mathbf{k}$ makes an angle $\beta$ with this axis as shown in Fig. 7.7. Let us again consider a right-handed photon with polarization vector $\mathbf{e}_{R}=\left(\mathbf{e}_{1}+i \mathbf{e}_{2}\right) / \sqrt{2}$. In the matrix element we then need the product

$$
\mathbf{e}_{\lambda}^{*} \cdot(\mathbf{k} \wedge \boldsymbol{\sigma})=\sqrt{\frac{1}{2}}\left[-\sin \beta(\mathbf{k} \wedge \boldsymbol{\sigma})_{z}+\cos \beta(\mathbf{k} \wedge \boldsymbol{\sigma})_{x}-i(\mathbf{k} \wedge \boldsymbol{\sigma})_{y}\right]
$$

which follows from (7.103). Calculating now the components of $\mathbf{k} \wedge \boldsymbol{\sigma}$ and expressing them in terms of the Pauli matrices $\sigma_{ \pm}$, one finds

$$
\begin{equation*}
\mathbf{e}_{\lambda}^{*} \cdot(\mathbf{k} \wedge \boldsymbol{\sigma})=i k \sqrt{\frac{1}{2}}\left[(1-\cos \beta) \sigma_{+}-(1+\cos \beta) \sigma_{-}+\sin \beta \sigma_{z}\right] \tag{8.78}
\end{equation*}
$$

after some algebra. This is essentially all we need in the following.
Let us now consider that the atom is initialy in the triplet state $T_{+}$with spin $m_{s}=+1$ along the $z$-axis. In the matrix element $\langle S| \mathbf{e}_{\lambda}^{*} \cdot(\mathbf{k} \wedge \boldsymbol{\sigma})\left|T_{+}\right\rangle$only the middle term in (8.78) will contribute. The whole calculation is then reduced down to the matrix element $S^{\dagger} \sigma_{-} T_{+}=-1 / \sqrt{2}$ we encountered above, giving

$$
\begin{equation*}
\langle S| \mathbf{e}_{\lambda}^{*} \cdot(\mathbf{k} \wedge \boldsymbol{\sigma})\left|T_{+}\right\rangle=i k \frac{1}{2}(1+\cos \beta) \tag{8.79}
\end{equation*}
$$

From (8.77) now follows the angular distribution

$$
\begin{equation*}
\frac{d \Gamma}{d \Omega}\left(T_{+} \rightarrow S+\gamma_{R}\right)=\frac{\alpha \omega}{16 \pi}\left(\frac{\hbar \omega}{m_{e} c^{2}}\right)^{2}(1+\cos \beta)^{2} \tag{8.80}
\end{equation*}
$$

for the emission of a positive-helicity photon.
When the photon is left-handed, the calculation is almost identical. The only change will be $\cos \beta \rightarrow-\cos \beta$ in the above result. If now this polarization is not measured, we must add these two contributions to the rate which then becomes

$$
\begin{equation*}
\frac{d \Gamma}{d \Omega}\left(T_{+} \rightarrow S+\gamma\right)=\frac{\alpha \omega}{8 \pi}\left(\frac{\hbar \omega}{m_{e} c^{2}}\right)^{2}\left(1+\cos ^{2} \beta\right) \tag{8.81}
\end{equation*}
$$

This is in agreement with the total rate (8.76) when we integrate over all directions of the photon. We would get the same rate if the atom was in one of the other intial states like $T_{0}$ or $T_{-}$.

One should notice that these angular distributions are esentally the same as for the corresponding electric transitions considered in Chapter 7. The underlying reason is that in both cases they take place in the dipole approximation from an initial atomic state with total angular momentum $J=1$ to a final state with $J=0$. They will therefore both have matrix elements proportional to the appropriate components of the corresponding rotation function $d_{m \lambda}^{(1)}(\beta)$ as previously mentioned.

## Chapter 9

## Many-electron atoms

The observed spectrum of the hydrogen atom is not much different from what follows from the solution of the Schrödinger equation in a pure Coulomb potential. Relativistic corrections are small since they are given by powers of the fine structure constant. And effects due to the spin of the electron are also small for the same reason.

In atoms with more than one electron, a new perturbation enters. This is the Coulomb interaction between the atoms. For the lightest atoms this is not small since it is of the same order as the interaction with the central nucleus. Standard perturbation theory will therefore not be very useful for these atoms. Also for heavier atoms the mutual interaction between the atoms play a central role. But when the number of electrons is suffiently large, we will see that it is possible to approximate all these non-central interactions with a new, effectice potential which will have central symmetry. Even if the actual values of atomic of the energy levels must now be found by numerical methods, the symmetry of the system allows for a straightforward classification of the different energy levels which will occur. Here the spins of the elctrons play a decisive role since the Pauli principle contrains the full wave function to be completely antisymmetric in all the variables describing the elctrons.

The important role of the electron spins we already see in the simplest many-electron atom, i.e. the helium atom with two electrons. Historically, this was one of the first important challenges of the new quantum mechanic which replaced the original, semiclassical ideas proposed by Bohr. We will therefore first look at it in some detail.

### 9.1 The helium atom

Let the two electrons in the neutral helium atom have positions $\mathbf{r}_{1}$ and $\mathbf{r}_{2}$ as shown in Fig. 9.1. Since the nucleus has the charge $Z=2$, the Hamiltonian becomes

$$
\begin{equation*}
H=-\frac{\hbar^{2}}{2 m}\left(\nabla_{1}^{2}+\nabla_{2}^{2}\right)-\frac{2 e^{2}}{4 \pi \epsilon_{0}}\left(\frac{1}{r_{1}}+\frac{1}{r_{2}}\right)+\frac{e^{2}}{4 \pi \epsilon_{0}} \frac{1}{\left|\mathbf{r}_{1}-\mathbf{r}_{2}\right|} \tag{9.1}
\end{equation*}
$$

where the last term represents the Coulomb repulsion between the two electrons. While the first two terms are centered around the nucleus, this term is non-central and depends on the separation $r_{12}=\left|\mathbf{r}_{1}-\mathbf{r}_{2}\right|$. Here lies the main complication of the problem, i.e.


Figure 9.1: Electron position coordinates in the helium atom.
finding the eigenvalues $E$ of the stationary Schrödinger equation $H \Psi(1,2)=E \Psi(1,2)$ for the corresponding eigenfunctions $\Psi(1,2)$.

The Hamiltonian can be written as $H=H_{0}+V$ where $H_{0}$ decribes two free electrons in a Coulomb potential while $V$ is the Coulomb repulsion between the electrons. If we now assume that this latter term is small, most of the properties of the atom will follow from the first and dominant term $H_{0}$. If we also ignore the symmetry properties of the full wave functions, they can simply be written as products $\Psi(1,2)=\psi_{a}\left(\mathbf{r}_{1}\right) \psi_{b}\left(\mathbf{r}_{2}\right)$ of hydrogen wave functions $\psi_{n \ell}(\mathbf{r})$ derived in Chapter 5 for a central charge $Z=2$. Here the first electron is in a state with quantum numbers denoted for simplicity by $a$ while the second has the quantum numbers denoted simply by $b$. The energy eigenvalues are therefore $E=\varepsilon_{a}+\varepsilon_{b}$ in this approximation, The energies of an unperturbed hydrogen atom are independent of the orbital angular momentum $\ell$ and thus we have for the helium atom energy levels in this approximation the result

$$
\begin{equation*}
E=-2^{2}\left(\frac{1}{n_{a}^{2}}+\frac{1}{n_{b}^{2}}\right) \mathrm{Ry} \tag{9.2}
\end{equation*}
$$

where $\mathrm{Ry}=e^{2} / 8 \pi \epsilon_{0} a_{0}=13.6 \mathrm{eV}$ is the (negative) ground state energy of the ordinary hydrogen atom with Bohr radius $a_{0}$.
In the lowest state of the helium atom both the electrons are in the hydrogen ground states $1 s$. Their product wave function is therefore $1 s 1 s$ which is usually written as $1 s^{2}$. It corresponds to the ground state energy $E_{0}=-4(1+1) \mathrm{Ry}=-8 \mathrm{Ry}=-108.8 \mathrm{eV}$. Comparing this with the measured value $E_{0}=-79.0 \mathrm{eV}$, we see that this lowest order decription is inaccurate to a much larger degree than what we encountered when discusing the hydrogen atom in previous chapters.

In the first excited state one of the electrons will be in the $2 s$ or $2 p$ state with the resulting excited energy $E_{1}=-4(1+1 / 4) R y=-5 R y$ for the helium atom. If this excited electron gets even more excited, it will eventually be liberated from the nucleus and we have an
unbound electron plus an ionized helium atom with energy $E_{\infty}=-4 \mathrm{Ry}=-54.4 \mathrm{eV}$. This implies that the helium atom in the state $2 s^{2}$, i.e. where both electrons are excited to the $2 s$ state, cannot be stable since it has an energy $E^{\prime}=-4(1 / 4+1 / 4) R y=-2 R y$ which is above that of an ionized atom. The state $2 s^{2}$ will thus decay into such an ionized state due to the internal dynamics of the atom when it is considered more accurately.

We must now include also the effect of the Coulomb repulsion between the electrons which we so far have assumed to be negligible. Since it tries to keep the electrons apart from each other in a classical description, its smallest value would correspond to the two electrons to be on the opposite side of the nucleus. In the unperturbed ground state $1 s^{2}$ of the atom, each electron is in a state with radius $a=a_{0} / 2$ and is therefore separated by the distance $r_{12}=a_{0}$. For this particular electron configuration the value of the repulsion energy is therefore $E_{e e}=e^{2} / 4 \pi \epsilon_{0}=2 \mathrm{Ry}=27.2 \mathrm{Ry}$. The ground state energy is therefore lifted up to $E_{0}=-6 \mathrm{Ry}=-81.6 \mathrm{eV}$ which is much closer to the experimental value. We therefore expect to see a similar improvement of the results when this term is included quantum mechanically.

Using perturbation there, the shift of the ground state energy will be given by the matrix element $\Delta E_{0}=\left\langle\Psi_{0}\right| \hat{V}\left|\Psi_{0}\right\rangle$ where the operator $\hat{V}$ is given by the above Coulomb repulsion. The ground state $\left|\Psi_{0}\right\rangle$ is given by the unperturbed wave function $1 s^{2}$ or more specifically as

$$
\begin{equation*}
\Psi_{0}(1,2)=\frac{1}{\pi}\left(\frac{2}{a_{0}}\right)^{3} e^{-2\left(r_{1}+r_{2}\right) / a_{0}} \tag{9.3}
\end{equation*}
$$

when we use the explicit form of the hydrogen wave function from Chapter 5. The result will therefore be given by the multiple integral

$$
\Delta E_{0}=\frac{e^{2}}{4 \pi^{3} \epsilon_{0}}\left(\frac{2}{a_{0}}\right)^{6} \int d^{3} r_{1} \int d^{3} r_{2} \frac{e^{-4\left(r_{1}+r_{2}\right) / a_{0}}}{\left|\mathbf{r}_{1}-\mathbf{r}_{2}\right|}
$$

which is highly non-trivial. But we can do it by a simple trick. From the result (6.57) for the Fourier transform of the Coulomb potential in Chapter 6, we know that it can be written as

$$
\frac{1}{4 \pi r}=\int \frac{d^{3} k}{(2 \pi)^{3}} \frac{e^{i \mathbf{k} \cdot \mathbf{r}}}{k^{2}}
$$

Using now this to get the denominator $r_{12}=\left|\mathbf{r}_{1}-\mathbf{r}_{2}\right|$ in (9.4) up into the exponent, it becomes

$$
\Delta E_{0}=\frac{e^{2}}{\pi^{2} \epsilon_{0}}\left(\frac{2}{a_{0}}\right)^{6} \int \frac{d^{3} k}{(2 \pi)^{3}} \frac{|I(k)|^{2}}{k^{2}}
$$

where the integral

$$
I(k)=\int d^{3} r e^{i \mathbf{k} \cdot \mathbf{r}-4 r / a_{0}}
$$

is exactly the same we met in (7.115) when we calculated the cross-section for photoproduction. From (7.116) we thus have

$$
I(k)=\frac{8 \pi\left(a_{0} / 4\right)^{3}}{\left[1+\left(k a_{0} / 4\right)^{2}\right]^{2}}
$$

and therefore for the Coulomb perturbation

$$
\Delta E_{0}=\frac{e^{2}}{\epsilon_{0}} \int \frac{d^{3} k}{(2 \pi)^{3}} \frac{1}{k^{2}\left[1+\left(k a_{0} / 4\right)^{2}\right]^{4}}
$$

Introducing $x=k a_{0} / 4$ as a new integration variable, this simplifies to

$$
\begin{equation*}
\Delta E_{0}=\frac{2 e^{2}}{\pi^{2} \epsilon_{0} a_{0}} \int_{0}^{\infty} \frac{d x}{\left(1+x^{2}\right)^{4}}=\frac{5 e^{2}}{16 \pi \epsilon_{0} a_{0}}=\frac{5}{2} \mathrm{Ry} \tag{9.4}
\end{equation*}
$$

where the value $5 \pi / 32$ of the last last integral must be looked up or done with contour integration in the complex plane. We thus get a perturbed ground state energy of $E_{0}=(-8+5 / 2) \mathrm{Ry}=-74.8 \mathrm{eV}$ which is a much better result. But we see that this first order perturbative correction is pretty large so higher order corrections will be needed to have a result we can trust. And these are very difficult to calculate - if they ever have.

A simpler and more accurate approach is to use the varionational method instead of perturbation theory. We then need only an approximative function with properties we would expect to find in the true wave function for the ground state. This construction will then involve unknown parameters which then have to be determined by minimizing the expectation value of the Hamiltonian (9.1).

We have already seen that the product wave function (9.3) describes the ground state with an accuracy of around $20 \%$. It describes two electrons moving freely around a central nuclear charge $Z=2$. But if we think a bit more physically about the situation we realize that this must be an approximation since from the point of view of one electron the electric field from the nucleus is reduced or screened by the other electron whenever this is 'inside' the first. A better approximation would then be to describe both electrons as still moving in a central Coulomb potential, but corresponding to a charge $Z<2$. In analogy with (9.3) we thus choose to work with the trial wave function

$$
\begin{equation*}
\Psi(1,2)=\frac{1}{\pi}\left(\frac{Z}{a_{0}}\right)^{3} e^{-Z\left(r_{1}+r_{2}\right) / a_{0}} \tag{9.5}
\end{equation*}
$$

where now $Z$ is the variational parameter.
Evaluation of the variational energy $\langle\Psi| \widehat{H}|\Psi\rangle$ for the Hamiltonian (9.1) is now straightforward. The first term gives twice the kinetic energy of an electron in the ground state of an hydrogen atom with nuclear charge $Z$ which becomes $2 Z^{2} e^{2} / 8 \pi \epsilon_{0} a_{0}$. The middle or central Coulomb term will for the same reason give the potential energy $-4 Z e^{2} / 4 \pi \epsilon_{0} a_{0}$ while the contribution for the last, repulsive Coulomb term follows from the above perturbative calculation. Comparing the two wave functions (9.5) and (9.3), we see that we just have to make the replacement $2 / a_{0} \rightarrow Z / a_{0}$ in the result (9.4) for the repulsive Coulomb energy, giving now $5 Z e^{2} / 32 \pi \epsilon_{0} a_{0}$. The total variational energy thus becomes

$$
\langle\Psi| \widehat{H}|\Psi\rangle=\frac{e^{2}}{4 \pi \epsilon_{0} a_{0}}\left(Z^{2}-4 Z+\frac{5}{8} Z\right)
$$

and is seen to have a minimum for $Z=Z_{\text {min }}=2-5 / 16=27 / 16$. This minimal value then follows as $E_{\text {min }}=-2 Z_{\text {min }}^{2} \mathrm{Ry}=-77.5 \mathrm{eV}$. It is surprisingly near the measured
value $E_{0}=-79.0 \mathrm{eV}$. It is also slightly above the true value, as it should be from the general considerations in Chapter 6 in connection with the variational principle.

The calculation of the ground state energy of the helium atom was iniziated by the Norwegian physicist Egil Hylleraas in 1928. He was a stipendiat with an American fellowship working with Max Born in Göttingen at a time when quantum mechanics was still very new and hotly debated. After a short time he obtained much more accurate results using the same varational principle but with a more general trial wave function than the simple version (9.5). This was based on the observation that the ground state must be rotational invariant and thus have total angular momentum $L=0$. It will therefore depend an fewer variables than the six components of the two position vectors $\mathbf{r}_{1}$ and $\mathbf{r}_{2}$ which in general would enter in an arbitrary wave function. Since the total potential energy in (9.1) is rotational invariant and only depends on the three lengths $r_{1}, r_{2}$ and $r_{12}$, these will then be the variables also for the ground state wave function. While $r_{1}$ and $r_{2}$ are obviously invariant under rotations, $r_{12}$ will also be since it depends on these two lengths and the angle $\theta$ between the corresponding vectors in Fig. 9.1. And this angle remains unchanged under a simultaneous rotation of both vectors.

His more general wave function was then taken to be the product of the exponential function (9.5) and a polynomial in these three variables. To leading order the improved trial wave function would then be

$$
\begin{equation*}
\Psi(1,2)=e^{-Z\left(r_{1}+r_{2}\right) / a_{0}}\left[C_{0}+C_{1} r_{12}+C_{2}\left(r_{1}+r_{2}\right)+C_{3}\left(r_{1}-r_{2}\right)^{2}+\ldots\right] \tag{9.6}
\end{equation*}
$$

For reasons given in the next section, we have here only included terms which are symmetric under the interchange of the two electrons. These extra terms decribes the exitation of the electrons due to the repulsive Coulomb force between them. In addition to $Z$, one now also has the coefficients $C_{n}$ as variational parameters. One of these is eliminated by normalization. With just these lowest order terms he could then obtain a result with an accuracy around one per mille. No one could at that time obtain more precise results. This work was later continued at the University of Oslo with his students where they also studied excited states of the helium atom and properties of other, similar atoms.

### 9.2 Electron spin and the Pauli principle

Accoring to the Pauli principle the full wave function $\Psi(1,2)$ of the two electrons must be antisymmetric on all the variables of the system, including the spin varibales. Since the Hamiltonian (9.1) only depends on the orbital variables $\mathbf{r}_{1}$ and $\mathbf{r}_{2}$, the full wave function can always be written as a direct product $\Psi(1,2)=\psi\left(\mathbf{r}_{1}, \mathbf{r}_{2}\right) \Phi\left(S, M_{S}\right)$. The total spin $S$ of the electrons can take the values $S=0$ and $S=1$ with corresponding $z$-components $M_{S}$. These states were found in the previous chapter. In particular, the singlet $S=0$ spin eigenfunction (8.62) we now write as

$$
\begin{equation*}
\Phi(0,0)=\sqrt{\frac{1}{2}}\left(\alpha_{1} \beta_{2}-\alpha_{2} \beta_{1}\right) \tag{9.7}
\end{equation*}
$$

It is antisymmetric in the interchange of the two labels 1 and 2 . The corresponding orbital wave function $\psi_{S}\left(\mathbf{r}_{1}, \mathbf{r}_{2}\right)$ must therefore be symmetric in the same two labels in
order to satisfy the Pauli principle. In perturbation theory ignoring for the moment the Coulomb repulsion between the electrons, one elctron will be in a hydrogenic state $a$ with the orbital wavefunction $\psi_{a}(\mathbf{r})$ and the other in the state $b$ with wavefunction $\psi_{b}(\mathbf{r})$. An allowed orbital wavefunction decribing both electrons is then

$$
\begin{equation*}
\psi_{S}\left(\mathbf{r}_{1}, \mathbf{r}_{2}\right)=\sqrt{\frac{1}{2}}\left[\psi_{a}\left(\mathbf{r}_{1}\right) \psi_{b}\left(\mathbf{r}_{2}\right)+\psi_{a}\left(\mathbf{r}_{2}\right) \psi_{b}\left(\mathbf{r}_{1}\right)\right] \tag{9.8}
\end{equation*}
$$

when they are in the singlet state with spin $S=0$. But in the triplet state with $S=1$, the spin eigenfunctions (8.61) are symmetric and the orbital wave function

$$
\begin{equation*}
\psi_{A}\left(\mathbf{r}_{1}, \mathbf{r}_{2}\right)=\sqrt{\frac{1}{2}}\left[\psi_{a}\left(\mathbf{r}_{1}\right) \psi_{b}\left(\mathbf{r}_{2}\right)-\psi_{a}\left(\mathbf{r}_{2}\right) \psi_{b}\left(\mathbf{r}_{1}\right)\right] \tag{9.9}
\end{equation*}
$$

is therefore antisymmetric in the two electron labels.
In the ground state both electrons must be in the lowest hydrogen wavefunction which we previously called $1 s$. We thus have the orbital configuration $[\mathrm{He}]=1 s^{2}$ for the two electrons in helium. But since the two quantum numbers $a$ and $b$ now are the same, we cannot form the antisymmetric combination (9.9). It vanishes. We can only form the symmetric wavefunction (9.8). The helium ground state must therefore be a spin singlet. And this we can say without any exact knowledge of the orbital wavefunction. Even inclusion of the Coulomb repulsion will not change this fact.

Since the orbital angular momentum $L$ of the ground state is also zero, the total momentum angular momentum $J$ is zero. Denoting an energy level of the atom with the quantum numbers $(J, L, S)$ by the symbol ${ }^{2 S+1} L_{J}$, the ground state is therefore given by ${ }^{1} S_{0}$. The $S$ here now stands for $L=0$ as previously used for the hydrogen atom. Including the Coulomb repulsion it will have the energy calculated in the previous section using the variational method.

Let us now consider the first radial excitation of this $L=0$ ground state. In the unperturbed picture one electron is then excited to the $2 s$ hydrogen orbital so that total electron configuration is $1 s 2 s$. Now we can form both a symmetric and an antisymmetric combination of these two orbitals and we have two resulting states for the atom, a spin singlet and spin triplet. Using the above notation, these two terms would then be denoted by ${ }^{1} S_{0}$ and ${ }^{3} S_{1}$. This ${ }^{1} S_{0}$ can then be said to be the first radial excitation of the ground state since it has the same quantum numbers. Needless to say, the new ${ }^{1} S_{0}$ state and the ${ }^{3} S_{1}$ will have the same energy as long as we ignore the Coulomb repulsion. But as for the ground state, it will increase their energies. In lowest order perturbation theory this shift is given by $\langle\Psi| \hat{V}|\Psi\rangle$ or

$$
\begin{equation*}
\Delta E=\int d^{3} r_{1} \int d^{3} r_{2} \psi^{*}\left(\mathbf{r}_{1}, \mathbf{r}_{2}\right) \frac{e^{2}}{4 \pi \epsilon_{0}\left|\mathbf{r}_{1}-\mathbf{r}_{2}\right|} \psi\left(\mathbf{r}_{1}, \mathbf{r}_{2}\right) \tag{9.10}
\end{equation*}
$$

The dominant contribution is seen to arise when the positions of the two electrons approach each other. But now we see that for the antisymmetric combination (9.9) the combined wave function approaches zero in this limit, while the symmetric combination (9.8) increases. This Coulomb shift of the triplet term ${ }^{3} S_{1}$ is therefore smaller than for the singlet ${ }^{1} S_{0}$. Thus, with the Coulomb repulsion, the three lowest terms or energy levels for the helium atom will be lifted and split up as seen in Fig.9.2.


Figure 9.2: Lowest energy level of the helium atom. The unperturbed levels on the left are lifted and split by the Coulomb repulsion between the electrons.

We have the surprising result that the electron spin plays an important role for the atomic energy levels even if doesn't explicitly appear in the Hamiltonian. This will be a very important effect also in atoms with more electrons as we soon will see.
Using the variational method one can also find the energies of these two excited states. For the ${ }^{3} S_{1}$ the trial wave function can be constructed along the same lines as was used to find the singlet function (9.6). We only have to ensure that it is now antisymmetric in the two electron coordinates. On the other hand, for the excited ${ }^{1} S_{0}$ state the wave function must be symmetric and also orthogonal to the ground state ${ }^{1} S_{0}$. This complicates the calculation. Instead one can use the Rayleigh-Ritz method outlined in Chapter 6 in a basis formed by the unperturbed $L=0$ hydrogen wave functions, but we will not pursue that here.

The first excited state with orbital angular momentum $L=1$ will similarly come from the unperturbed configuration $1 s 2 p$. Again including the Coulomb repulsion, this is split up in a spin triplet of states below a singlet state. This latter will have total angular momentum $J=1$ and thus be denoted by ${ }^{1} P_{1}$. Since the triplet state has electron spin $S=1$, the total angular momentum will then take the values $J=2,1,0$. We thus find the three triplet terms ${ }^{3} P_{2},{ }^{3} P_{1}$ and ${ }^{3} P_{0}$. Including the relativistic spin-orbit coupling as in the previous chapter, these will be split slightly apart.

### 9.3 Central-field approximation

A heavier atom with $N$ electrons moving in the Coulomb field of a nucleus with charge $Z e$ has the Hamiltonian

$$
\begin{equation*}
H=\sum_{i=1}^{N}\left(-\frac{\hbar^{2}}{2 m} \nabla_{i}^{2}-\frac{Z e^{2}}{4 \pi \epsilon_{0} r_{i}}\right)+\sum_{i<j} \frac{e^{2}}{4 \pi \epsilon_{0}} \frac{1}{\left|\mathbf{r}_{i}-\mathbf{r}_{j}\right|} \tag{9.11}
\end{equation*}
$$

The physical states are eigenfunctions $\Psi(1,2, \ldots, N)$ which are completely antisymmetric in all the electron variables. Even if the problem then has this simple formulation, the solution is mathematically very difficult and must be done numerically. But given that, it explains most properties of the atoms and thus the matter we see around us. This is the realm of chemists.

Each electron has an angular momentum operator $\mathbf{L}_{k}=-i \hbar \mathbf{r}_{k} \wedge \nabla_{k}$. Each of these are conserved by the first part of the above Hamiltonian because it is central as for the ordinary hydrogen atom. But $\mathbf{L}_{k}$ does not commute with the last term since

$$
\begin{aligned}
{\left[\mathbf{L}_{k}, \frac{1}{r_{i j}}\right] } & =\left(\mathbf{L}_{k} \frac{1}{r_{i j}}\right)=-i \hbar \mathbf{r}_{k} \wedge\left(\nabla_{k} \frac{1}{r_{i j}}\right) \\
& =-i \hbar\left[-\mathbf{r}_{k} \wedge \frac{\mathbf{r}_{i}-\mathbf{r}_{j}}{r_{i j}^{3}} \delta_{i k}+\mathbf{r}_{k} \wedge \frac{\mathbf{r}_{i}-\mathbf{r}_{j}}{r_{i j}^{3}} \delta_{j k}\right]
\end{aligned}
$$

is non-zero. We can therefore not label the general solutions by the angular momenta of the individual electrons. But for the total angular momentum

$$
\begin{equation*}
\mathbf{L}=\sum_{k=1}^{N} \mathbf{L}_{k} \tag{9.12}
\end{equation*}
$$

the commutator is zero as follows from

$$
\left[\mathbf{L}, \frac{1}{r_{i j}}\right]=-i \hbar \sum_{k=1}^{N}\left[-\mathbf{r}_{k} \wedge \frac{\mathbf{r}_{i}-\mathbf{r}_{j}}{r_{i j}^{3}} \delta_{i k}+\mathbf{r}_{k} \wedge \frac{\mathbf{r}_{i}-\mathbf{r}_{j}}{r_{i j}^{3}} \delta_{j k}\right]=0
$$

since the vector wedge product is antisymmetric. Thus the full states can be labelled by the total angular momentum quantum number $L$ and the $z$-component $M_{L}=L_{z} / \hbar$.

From the look of the Hamiltonian one would also think that the spin $\mathbf{S}_{k}=(\hbar / 2) \boldsymbol{\sigma}_{k}$ of each electron also is a good quantum number since it doesn't appear at all. But then we are deceived. Imposing the Pauli principle restriction on the total wave functions we know from the previous discussion of the helium atom that the effect of the Coulomb repulsion depends on the total spin angular momenta $\mathbf{S}$ defined similarly to the orbital part $\mathbf{L}$ in (9.12). Thus the total spin angular momentum quantum number $S$ and the $z$-component $M_{S}=S_{z} / \hbar$ will also be important quantum numbers for the physical states.

While the physics of the first part of the Hamiltonian (9.11) describes non-interacting electrons moving in the central Coulomb potential from the nucleus, the last term is due to the mutual Coulomb repulsion between the electrons. As for the helium atom, the main complications of the problem lies here. But with many electrons this term
is even more important in determining the physical properties of atoms. The reason is that the number of such terms grows like $N^{2} / 2$ when $N$ is large and thus can become of the same magnitude as the central potential in the first term. It can thus not be treated with perturbation theory which worked quite well for the helium atom.

But when the number $N-1$ of electrons each electron interacts with, is large, we are allowed to consider the effect of all these interactions as a central potential $U(r)$ acting on this particular electron plus a remaining, non-central term. More precisely, we can then write

$$
\begin{equation*}
\sum_{i<j} \frac{e^{2}}{4 \pi \epsilon_{0}} \frac{1}{\left|\mathbf{r}_{i}-\mathbf{r}_{j}\right|}=\sum_{i=1}^{N} U\left(r_{i}\right)+H^{\prime} \tag{9.13}
\end{equation*}
$$

where $H^{\prime}$ now is non-central, involving all the electron coordinates. Including relativistic effects, there is also a spin-orbit coupling which also can be treated within perturbation theory. We will come back to that later in the chapter.
The full Hamiltonian (9.11) can then be written as $H=H_{0}+H^{\prime}$ where

$$
\begin{equation*}
H_{0}=\sum_{i=1}^{N}\left[-\frac{\hbar^{2}}{2 m} \nabla_{i}^{2}-\frac{Z e^{2}}{4 \pi \epsilon_{0} r_{i}}+U\left(r_{i}\right)\right] \tag{9.14}
\end{equation*}
$$

now decribes non-interacting electrons moving in a common, effective potential

$$
\begin{equation*}
V(r)=-\frac{Z e^{2}}{4 \pi \epsilon_{0} r}+U(r) \tag{9.15}
\end{equation*}
$$

And the eigenvalue problem

$$
\begin{equation*}
\left[-\frac{\hbar^{2}}{2 m} \boldsymbol{\nabla}^{2}+V(r)\right] \psi_{a}(\mathbf{r})=\varepsilon_{a} \psi_{a}(\mathbf{r}) \tag{9.16}
\end{equation*}
$$

we assume we can solve as accurately as we want, at least numerically.
For short distances $r \rightarrow 0$ the electron sees the full charge of the nucleus and the effective potential must approach the classical value $V(r) \rightarrow-Z e^{2} / 4 \pi \epsilon_{0} r$. In the opposite limit $r \rightarrow \infty$ the electron is outside the remaining $N-1$ electrons. It thus sees the nucleus with charge $Z e$ reduced by the amount $(N-1) e$ and the effective potential should approach $V(r) \rightarrow-(Z-N+1) e^{2} / 4 \pi \epsilon_{0} r$. A neutral atom has $N=Z$ and the screening is correspondingly strong.

Since the effective potential (9.15) is not a pure Coulomb potential, the energy levels will also depend on the angular momentum quantum number $\ell$ as discussed in Chapter 5. The principal quantum number $n$ is for a non-Coulombic potential defined by the number of nodes is the radial part of the wavefunction as for the hydrogen atom. Each value of $n$ determines an electronic shell. The $K$-shell electrons have $n=1$, in the $L$-shell they have $n=2$ and so on. These two quantum numbers are here and in the following denoted by Latin lower case letters like $a$ and $b$. These one-particle solutions are often called orbitals or subshells. Electrons in higher shells will be seen to have overlapping energy values.
We don't know yet how to find the general shape of the effective potential. But it will turn out that it can be calculated and found to give a resulting electron interaction
$H^{\prime}$ which is so small that it can be used in perturbation theory based on the solutions of the much simpler problem defined by $H_{0}$. And these non-interacting many-electron solutions are easily constructed in terms of the single-particle orbitals in the same way as for helium taking into account the Pauli principle. We get the idea when we write the helium ground state $\Psi(1,2)$ as the product between the orbital part $1 s^{2}$ and the spin singlet (9.7) as

$$
\begin{aligned}
\Psi(1,2) & =\psi_{1 s}\left(\mathbf{r}_{1}\right) \psi_{1 s}\left(\mathbf{r}_{2}\right) \Phi(0,0)=\sqrt{\frac{1}{2}} \psi_{1 s}\left(\mathbf{r}_{1}\right) \psi_{1 s}\left(\mathbf{r}_{2}\right)\left(\alpha_{1} \beta_{2}-\alpha_{2} \beta_{1}\right) \\
& =\sqrt{\frac{1}{2}}\left|\begin{array}{ll}
\psi_{1 s}\left(\mathbf{r}_{1}\right) \alpha_{1} & \psi_{1 s}\left(\mathbf{r}_{1}\right) \beta_{1} \\
\psi_{1 s}\left(\mathbf{r}_{2}\right) \alpha_{2} & \psi_{1 s}\left(\mathbf{r}_{2}\right) \beta_{2}
\end{array}\right|
\end{aligned}
$$

The resulting form is called a Slater determinant. Under interchange of the two electron labels, the antisymmetry of the full wavefunction follows directly from the antisymmetry of the determinant under the interchange of two rows.
For more than two electrons we then know how to proceed. Denoting the full orbitals by $\psi_{A}(i) \equiv \psi_{a}\left(\mathbf{r}_{i}\right) \chi_{i}$ where $\chi$ is one of the eigenspinors $\alpha$ or $\beta$, the full wavefunction satisfying the Pauli principle will be given by the Slater determinant

$$
\Psi(1,2, \ldots, N)=\sqrt{\frac{1}{N!}}\left|\begin{array}{cccc}
\psi_{A}(1) & \psi_{B}(1) & \ldots & \psi_{F}(1)  \tag{9.17}\\
\psi_{A}(2) & \psi_{B}(2) & \ldots & \psi_{F}(2) \\
\vdots & \vdots & \ddots & \vdots \\
\psi_{A}(N) & \psi_{B}(N) & \ldots & \psi_{F}(N)
\end{array}\right|
$$

where the labels $A, B, \ldots, F$ depend on which orbitals the electrons occupy. For instance, with two electrons and total electron spin component $M_{S}=+1$, we get the state

$$
\sqrt{\frac{1}{2}}\left|\begin{array}{ll}
\psi_{a}\left(\mathbf{r}_{1}\right) \alpha_{1} & \psi_{b}\left(\mathbf{r}_{1}\right) \alpha_{1} \\
\psi_{a}\left(\mathbf{r}_{2}\right) \alpha_{2} & \psi_{b}\left(\mathbf{r}_{2}\right) \alpha_{2}
\end{array}\right|=\sqrt{\frac{1}{2}}\left[\psi_{a}\left(\mathbf{r}_{1}\right) \psi_{b}\left(\mathbf{r}_{2}\right)-\psi_{a}\left(\mathbf{r}_{2}\right) \psi_{b}\left(\mathbf{r}_{1}\right)\right] \alpha_{1} \alpha_{2}
$$

which is a general spin triplet with an antisymmetric orbital part (9.9) as already used for helium. Had we constructed the corresponding state with $M_{S}=0$, we would have found a superposition of a spin singlet and a spin triplet. This just illustrates that the Slater determinant does not in general give states which are angular momentum or spin eigenstates.

For this approximation to be useful, we must determine the effective potential (9.15). There are several methods available and one is outlined in the last section of this chapter. Given this, we solve the one-particle Schrödinger equation (9.16) which determines the energy $\varepsilon_{a}$ of an electron in the orbital with wavefunction $\psi_{a}(\mathbf{r})$ with quantum numbers $a=(n \ell)$. The dependence of the eigenvalues on the orbital angular momentum quantum number $\ell$ has already been explained due to the non-Coulombic form of the potential. Since the effective potential depends on the atomic number $Z$, these eigenvalues will also have a dependence on this parameter. A representative diagram of the lowest orbital energies are shown in Fig.9.3. Orbitals in the same shell with lower values of $\ell$ will be closer to the nucleus and thus less screened. Their energies are therefore reduced or pulled down in the diagram compared with those with larger values of this quantum


Figure 9.3: Schematic orbital energy levels for a general atom. On the right the eigenvalues are projected on the same axis in order to see the ordering of orbitals.
number. One finds that the energies of the $3 d$ and $4 s$ orbitals are very close, and their individual ordering depends on the atomic atom $Z$. The same applies to the orbitals $4 d$ and $5 s$.

Using now the Pauli principle, we can build up the periodic table in this central field approximation. We have already considered helium with two electrons has the electronic configuration $[\mathrm{He}]=1 s^{2}$. Lithium has one more electron which is seen from the figure to enter the $L$-shell in the 2 s orbital, i.e. $[\mathrm{Li}]=1 s^{2} 2 s$. Beryllium has four electrons and therefore $[\mathrm{Be}]=1 s^{2} 2 s^{2}$. In a more compact notation we can write this as $[\mathrm{Be}]=[\mathrm{He}] 2 s^{2}$ which will especially be useful for heavier atoms.

These two subshells are now filled with electrons. When we next come to boron with five electrons, this new electron must enter the $2 p$ orbital and thus has the configuration $[\mathrm{B}]=[\mathrm{He}] 2 s^{2} 2 p$. This orbital can occupate six electrons and when it is filled, we have the neon atom $[\mathrm{Ne}]=[\mathrm{He}] 2 s^{2} 2 p^{6}$. Like helium this is a forms noble gas since the neon atom is very stable due to the relative high energy needed to change it, corresponding to the energy gap between the $2 p$ and $3 s$ orbitals in the figure.

Filling now the next two orbitals $3 s$ and $3 p$, we construct the elements starting with sodium $[\mathrm{Na}]=[\mathrm{Ne}] 3 s$ and ending with argon $[\mathrm{Ar}]=[\mathrm{Ne}] 3 s^{2} 3 p^{6}$. It forms a noble gas for the same reason as neon does. An extra electron will now go into the $4 s$ orbital because for $Z$ around twenty the $3 d$ orbital is slightly higher in energy. We then have potassium $[\mathrm{K}]=[\mathrm{Ar}] 4 s$. After calcium $[\mathrm{Ca}]=[\mathrm{Ar}] 4 s^{2}$, the $3 d$ orbital starts to fill up. Since these
have a smaller principal quantum number, they will be on the inside of the $4 s$-electrons. The $3 d$-orbital can accomodate at most ten electrons. Filling it up, we generate ten elements. All will have two outer electrons in the $4 s$ orbital which make them metallic and with similar chemical properties. For instance, among these transition elements there is iron $[\mathrm{Fe}]=[\mathrm{Ar}] 3 d^{6} 4 s^{2}$ and finally zinc $[\mathrm{Zn}]=[\mathrm{Ar}] 3 d^{10} 4 s^{2}$.

The next electrons goes into the $4 p$ orbital. When it is filled with six electrons, we have the noble gas krypton $[\mathrm{Kr}]=[\mathrm{Ar}] 3 d^{10} 4 s^{2} 4 p^{6}$. With even more electrons, we first fill the $5 s$-orbital which now is slightly below the $4 d$ orbital in energy, but is outside it in space. When next the inner $4 d$ orbital is filled up, we create a new transition series of elements with similar chemical properties. It ends with cadmium $[\mathrm{Cd}]=[\mathrm{Kr}] 4 d^{10} 5 s^{2}$ which has atomic number $Z=48$. Knowing the energies of the higher orbitals, we can continue in exactly the same way and build up the remaining part of periodic table.

### 9.4 Terms and the Hund rules

In the central-field approximation where we neglect the remaining Coumomb repulsion $H^{\prime}$ defined by (9.13), there will be many electronic configurations constructed from different Slater determinants that have the same energy. They are in general states without spesific values of the angular momentum quantum numbers $L$ and $S$ as already explained. The effect of the perturbation $H^{\prime}$ must therefore be evaluated using degenerate perturbation theory discussed in Chapter 6. As pointed out there, it is then very convenient to combine the degenerate eigenstates of $H_{0}$ into a new basis in which the matrix elements of $H^{\prime}$ are diagonal. And since we have already shown that $H^{\prime}$ commutes with the operators $\mathbf{S}$ and $\mathbf{L}$, we must out of the degenerate electronic configurations form linear combinations which are eigenstates of these operators, i.e. with definite values for the total electronic spin $S$ and angular momentum $L$.

A filled orbital will always have $S=0$ and $L=0$. This follows directly from the Pauli principle which says that for every electron spin pointing up, there will be one pointing down in the same orbital. The same applies to the orbital angular momentum. So we only have to consider non-filled orbitals.

It is simplest to consider first two electrons in different shells, but let's say both in $p$-orbitals and thus with $\ell=1$. The electrons are then said to be inequivalent. Together they can then form states with total angular momentum $L=0,1,2$. Since they also can be grouped into combinations with total spin $S=0$ and $S=1$, we obtain three angular momentum singlet eigenstates ${ }^{1} S,{ }^{1} P$ and ${ }^{1} D$ and the three corresponding triplet states ${ }^{3} S,{ }^{3} P$ and ${ }^{3} D$. In this basis the Coulomb perturbation $H^{\prime}$ will be diagonal. As for helium, the electrons in the triplet states are wider apart from each other and will thus have lower energy than the singlet states. Moreover, the $D$-states are also wider apart from each other than the electrons in the $P$-states and will therefore be lower in energy. For the very same reason the $P$-states will be lower in energy than the $S$-states. This is schematically illustrated in Fig.9.4.

The three lowest terms in lowest energy state ${ }^{3} D_{1,2,3}$ are still degenerate. This degeneracy is lifted when we include the spin-orbit couplings $\sum_{k} f_{k} \mathbf{L}_{k} \cdot \mathbf{S}_{k}$ where the functions $f_{k}$ involve only radial coordinates. With help of the Wigner-Eckart theorem discussed


Figure 9.4: The degenerate electron configuration $n p n^{\prime} p$ is split apart by the Coulomb repulsion into terms with definite values of spin and angular momentum.
in the previous chapter, this interaction is equivalent to a perturbation proportional to the operator $\mathbf{L} \cdot \mathbf{S}$. As for the fine-structure in the hydrogen atom, it splits the level into terms with definite values for the total angular momentum $\mathbf{J}=\mathbf{L}+\mathbf{S}$. A more detailed look at this perturbation shows that it is the term with the lowest value of $J$ which has the lowest energy. The ground state is therefore the term ${ }^{3} D_{1}$ and the higher terms will be excited states of the atom.

When the two electrons are in the same shell, they are said to be equivalent. They then have the the same value for the principal quantum number $n$ and many of the above terms will no longer be allowed because of the Pauli principle. This applies for instance to the two $p$ electrons in the carbon atom $[\mathrm{C}]=[\mathrm{He}] 2 s^{2} 2 p^{2}$. The two electrons in the $2 s$-orbital form a closed subshell and can therefore be ignored here. When one now constructs the possible angular momentum eigenstates, one finds that only the terms ${ }^{1} S_{0},{ }^{1} D_{2}$ and ${ }^{3} P_{0,1,2}$ will appear where the degenerate spin triplet is lowest in energy as expected for the same reasons as before. It will be split apart by the spin-orbit coupling so that the $J=0$ term gets the lowest energy as shown in Fig.9.5. The ground state of the neutral carbon atom is therefore ${ }^{3} P_{0}$ and thus has no total spin. This will be somewhat different when the atom binds to other atoms. The situation then becomes a bit more intricate - and interesting for organic chemists.
The net effect of all these interactions to determine the physical ground state of an atom with several equivalent electrons, can be summed up in the three Hund rules. They are


Figure 9.5: Term digram for the neutral carbon atom with two equivalent $p$ electrons.
not absolutely true, but emperically established and supported by detailed calculations. They give the following step-by-step procedure to follow in this construction:

1. The term with largest possible value of the total spin $S$ has the lowest energy.
2. For a given value of $S$, the term with largest value of the orbital angular momentum $L$ has the lowest energy.
3. The term with the smallest value $J=|L-S|$ of the total angular momentum is then the ground state when the orbital is less than half-filled, if not it is the term with the maximum value $J=L+S$.

The first rule follows from the already mentioned fact that the terms with largest value of $S$, has the most orbital antisymmetry and therefore contains electrons wthat prefer to stay away from each other and thus reduce the effect of the Coulomb repulsion. The same applies to terms with the largest value of $L$ as in the second rule, while the third rule is direct consequence of the spin-orbit coupling $\mathbf{L} \cdot \mathbf{S}$. For given values of $L$ and $S$, it takes on a extremal value for $J=|L \pm S|$. More detailed calculations show that it has a prefactor that is positive when the subshell is less than half-filled and negative otherwise. Thus $J=|L-S|$ for the term with lowest energy in the first case.
There is a simple way to systematically determine the quantum numbers of the ground state involving equivalent electrons. We consider 'boxes' for each value of the $z$-component
$m_{\ell}$ of the orbital angular momentum. There will then be totally $2 \ell+1$ such boxes available for electrons in the orbital. To satisfy the Pauli principle, two electrons in the same box must then have opposite spin directions with $z$-component $m_{s}= \pm 1 / 2$.

This is illustrated in Fig.9.6 for a p-orbital, i.e. with three boxes. We want to place two electrons in the orbital as for the carbon atom. One possibility suggested by the first Hund rule, is shown to the left. If it was allowed, it would have a total orbital quantum number $M_{L}=1+1=2$ and thus give a $D$ term as occurs in Fig.9.4 with inequivalent electrons. But now the Pauli principle forbids it since two electrons cannot have all quantum numbers the same. Remember that equivalent electrons have the


Figure 9.6: Two equivalent electrons in a $p$-orbital. While the assignment in a) is forbidden because of the Pauli principle, the assignment in b) is allowed and suggested by the two first Hund rules.
same principal quantum number. The term with maximum number of $M_{S}$ and thus of $S$ will be formed by having a spin-up electron in two different boxes. It thus has $M_{S}=1 / 2+1 / 2=1$ and therefore $S=1$. Next we use the second rule and place the two electrons as shown to the right in the figure, i.e. with $M_{L}=0+1=1$ and thus $L=1$. The first and second Hund rules thus gives a ${ }^{3} P_{0,1,2}$ term as already said above. Since a $p$-orbital with only two electrons, is less than half-filled, the spin-orbit coupling now makes the $J=0$ term ${ }^{3} P_{0}$ lowest in energy and thus the ground state. Needless to say, this agrees with what concluded above.

A perhaps more interesting example is the iron atom with the electron configuration $[\mathrm{Fe}]=[\mathrm{Ar}] 3 d^{6} 4 s^{2}$. The $4 s$-orbital is full and will thus not contribute any spin or angular momentum. There are in addition six electrons in the $3 d$-orbital. It can take at most ten electrons. To maximize the total spin $S$, we must have five electrons with spin up and therefore one with spin down. One box is therefore doubly occupied while the four other boxes contain single electrons with spins in the same directions. For the total $z$-component of the spin we then have $M_{S}=2$ and therefore also $S=2$. Using now the second Hund rule, the doubly occupied box must have a maximum value of the quantum number $\left|m_{\ell}\right|$, let's say the two electrons are in the box with $m_{\ell}=+2$ as in Fig.9.7. The orbital momenta thus add up to give $M_{L}=(-2)+(-1)+0+(+1)+2(+2)=2$ and therefore $L=2$. Since the orbital is more than half-filled, the total spin angular momentum is thus $J=2+2=4$ according to the third Hund rule. In this way we find the ground state of iron to be the term ${ }^{5} D_{4}$. It thus has a relatively large spin.


Figure 9.7: Six equivalent electrons in a $3 d$-orbital of the iron atom forming the ground state.

### 9.5 Atomic magnetic moments

When the atom is placed in an external magnetic field $\mathbf{B}$, the electrons will couple to it as described in Chapter 7. The total interaction is therefore

$$
\begin{equation*}
H_{B}=-\frac{e}{2 m_{e}} \sum_{k=1}^{N}\left(\mathbf{L}_{k}+2 \mathbf{S}_{k}\right) \cdot \mathbf{B} \tag{9.18}
\end{equation*}
$$

where $m_{e}$ is the electron mass. If the field is along the $z$-axis, the perturbation becomes

$$
\begin{equation*}
H_{B}=-\frac{e}{2 m_{e}} \sum_{k=1}^{N}\left(L_{k z}+2 S_{k z}\right)=-\frac{e}{2 m_{e}}\left(J_{z}+S_{z}\right) \tag{9.19}
\end{equation*}
$$

Only electrons in non-filled subshells will contribute. When the terms are split apart by the spin-orbit coupling, the final states are eigenstates of the total angular momentum $\mathbf{J}$. In the basis formed by these states, we calculate the matrix elements of this perturbation as we did for the hydrogen atom in the previous chapter. We then showed that the results could be described by the introduction of an effective, magnetic moment operator (8.44)

$$
\begin{equation*}
\boldsymbol{\mu}_{e f f}=\frac{e g}{2 m_{e}} \mathbf{J} \tag{9.20}
\end{equation*}
$$

where $g$ is the Landé $g$-factor (8.48). In our case it will involve the total spin and angular momenta of the atom, i.e.

$$
\begin{equation*}
g=1+\frac{J(J+1)+S(S+1)-L(L+1)}{2 J(J+1)} \tag{9.21}
\end{equation*}
$$

It determines the absolute magnitude of the magnetic moment.
The effective coupling of the atom to the field is now $H_{e f f}=-\boldsymbol{\mu}_{e f f} \cdot \mathbf{B}$ and has a value proportional with the $z$-component $J_{z}$. It is conventional to define the magnitude $\mu_{J}$ of the magnetic moment by the value it takes when $\mathbf{J}$ is along the magnetic field, i.e. by the maximal value $J_{z}=J \hbar$. It thus becomes

$$
\begin{equation*}
\mu_{J}=\frac{e \hbar}{2 m_{e}} g J=\mu_{B} g J \tag{9.22}
\end{equation*}
$$

where $\mu_{B}$ is the Bohr magneton.
With the help of Hund's rules, we can now determine the quantum numbers $L, S, J$ of the ground state of the atom. Thus we also can calculate directly the value of its
magnetic moment. For instance, for iron considered in the previous section, we found $S=2, L=2$ and $J=4$. Inserted into the Lané formula (9.21), we get $g=3 / 2$ and thus the magnetic moment $\mu_{J}=\mu_{B} \cdot(3 / 2) \cdot 4=6 \mu_{B}$. This is a rather large value and shows that iron has strong magnetic properties.

An other example is vanadium with the electronic configuration $[\mathrm{V}]=[\mathrm{Ar}] 3 d^{3} 4 s^{2}$. Placing the three $d$-electrons in boxes as dictated by Hund's rules, the spin angular moementa


Figure 9.8: Three equivalent $d$-electrons defining the ground state of the vanadium atom.
for the ground state can be read off Fig.9.8. We see that the total spin $S=3 / 2$ and $L=3$. Since the orbital is less than half-filled, we will have $J=3-3 / 2=3 / 2$. Thus we have the ground state ${ }^{4} F_{3 / 2}$. The Landé factor becomes $g=2 / 5$ giving a magnetic moment $\mu_{J}=(3 / 5) \mu_{B}$. It should thus be less magnetic than iron.

But the connection between these atomic magnetic moments and the magnetism observed in metals is not so direct. First, the atomic orbitals used here will be distorted by neighbouring atoms. Secondly, the metallic binding involves electrons which can move around almost freely between different atoms and will also contribute to the magnetic properties of the metal. A complete understanding of all these different magnetic effects are left for a more advanced course in solid state physics.

### 9.6 Hartree approximations

So far we have not said anything about how one systematically can determine the central potential $U(r)$ defined in (9.13) and caused by the screening of other electrons. At a position $\mathbf{r}^{\prime}$ they will create a charge density

$$
\rho\left(\mathbf{r}^{\prime}\right)=-e \sum_{b \neq a}\left|\psi_{b}\left(\mathbf{r}^{\prime}\right)\right|^{2}
$$

where the quantum numbers $a$ describe the orbital of the electron being acted upon. Here we should also include the electron spin and therefore use the full quantum numbers introduced earlier, $A=(a \sigma)$ with $\sigma=\uparrow$ for spin up and $\sigma=\downarrow$ for spin down. But to simplify matters a little, we drop that and thus consider electrons to be spinless fermions.

The potential energy of the electron at position $\mathbf{r}$ in the orbital $\psi_{a}$ created by this charge distribution is now

$$
\begin{equation*}
U(\mathbf{r})=-e \int d^{3} r^{\prime} \frac{\rho\left(\mathbf{r}^{\prime}\right)}{4 \pi \epsilon_{0}\left|\mathbf{r}-\mathbf{r}^{\prime}\right|}=\frac{e^{2}}{4 \pi \epsilon_{0}} \sum_{b \neq a} \int d^{3} r^{\prime} \frac{\psi_{b}^{*}\left(\mathbf{r}^{\prime}\right) \psi_{b}\left(\mathbf{r}^{\prime}\right)}{\left|\mathbf{r}-\mathbf{r}^{\prime}\right|} \tag{9.23}
\end{equation*}
$$

This should now be used to form the effective potential (9.15). However, we now see at least two problems with this approach. First, the potential is not in general radially symmetric, but depends on all three components of the position vector $\mathbf{r}$. This we can fix by taking some angular average. Secondly, we see that the potential depends on which orbital $\psi_{a}$ we consider. We will find the effective potential being in general different for different orbitals. Again we can get out of that problem by taking some average in constructing a general potential, but this dependence is really inherent to the method.
Anyhow, the resulting Schrödinger equation (9.16) now takes the form

$$
\begin{equation*}
\left(-\frac{\hbar^{2}}{2 m} \nabla^{2}-\frac{Z e^{2}}{4 \pi \epsilon_{0} r}\right) \psi_{a}(\mathbf{r})+\frac{e^{2}}{4 \pi \epsilon_{0}} \sum_{b \neq a} \int d^{3} r^{\prime} \frac{\psi_{b}^{*}\left(\mathbf{r}^{\prime}\right) \psi_{b}\left(\mathbf{r}^{\prime}\right)}{\left|\mathbf{r}-\mathbf{r}^{\prime}\right|} \psi_{a}(\mathbf{r})=\varepsilon_{a} \psi_{a}(\mathbf{r}) \tag{9.24}
\end{equation*}
$$

and would give the orbital energies $\varepsilon_{a}$. But the equation is really a set of coupled, non-linear differential equations, first proposed by Hartree. They can only be solved numerically and then iteratively. Here it would mean that we would intially assume some orbital functions $\psi_{b}\left(\mathbf{r}^{\prime}\right)$ similar to ordinary hydrogenic eigenfunctions. These would then be used to calculate the potential (9.23). With this achieved, one then solves the Hartree equation (9.24) and finds the new orbitals $\psi_{a}$. These should then decribe the interacting electrons better. Then one goes back to (9.23) with the improved orbitals and repeats the process. Hopefully the iteration converges so that one ends up with stable results for the orbitals and their energies.

But in this method there has so far been no worry about the electrons being fermions satisfying the Pauli principle. This was first done by V. Fock with the help of Slater determinants for the electron configuration. The net result is not so much different from the Hartree equation (9.24), but involves an extra term,

$$
\begin{align*}
\left(-\frac{\hbar^{2}}{2 m} \nabla^{2}-\frac{Z e^{2}}{4 \pi \epsilon_{0} r}\right) \psi_{a}(\mathbf{r}) & +\frac{e^{2}}{4 \pi \epsilon_{0}} \sum_{b \neq a} \int d^{3} r^{\prime} \frac{\psi_{b}^{*}\left(\mathbf{r}^{\prime}\right) \psi_{b}\left(\mathbf{r}^{\prime}\right)}{\left|\mathbf{r}-\mathbf{r}^{\prime}\right|} \psi_{a}(\mathbf{r})  \tag{9.25}\\
& -\frac{e^{2}}{4 \pi \epsilon_{0}} \sum_{b \neq a} \int d^{3} r^{\prime} \frac{\psi_{b}^{*}\left(\mathbf{r}^{\prime}\right) \psi_{a}\left(\mathbf{r}^{\prime}\right)}{\left|\mathbf{r}-\mathbf{r}^{\prime}\right|} \psi_{b}(\mathbf{r})=\varepsilon_{a} \psi_{a}(\mathbf{r})
\end{align*}
$$

This is called the exchange term and is always important in many-fermion systems. Since the electron at the position $\mathbf{r}$ in the orbital $\psi_{a}$ is identical with the electron at $\mathbf{r}^{\prime}$ in the orbital $\psi_{b}$ setting up the potential, they must enter the full interaction in an antisymmetric way. With this extra term we can now also drop the requirement $b \neq a$ in the summations because now the contribution to the sums from the terms with $b=a$ will cancel between them. Thus with the Hartree-Fock approximation we have the possibility to obtain an effective potential valid for all the orbitals and thus more physical. But notice that the exchange term can no longer be understood as a classical potential. It is a pure quantum contribution.

## Chapter 10

## Non-relativistic quantum fields

Quantum mechanics of systems with many interacting particles becomes quite cumbersome when their number becomes large as for heavy atoms, but also in nuclear physics and condensed matter physics. With fermions we have to construct physical states out of very large Slater determinants or corresponding symmetric states for bosons. However, since the particles are identical, it is possible to simplify the mathematics enormously by considering them as quanta of an underlying field. The particles will then be like quanta of the harmonic oscillator or photons of the electromagnetic field.
We will only consider non-relativistic motion of these particles. The corresponding field theory is then also non-relativistic and often called Schrödinger field theory. At the end we will return to the problem of a particles on a periodic lattice which we also discussed at the end of Chapter 3. Including now interactions, we can then make contact with modern problems in solid state physics.

### 10.1 Identical particles

Let us consider $N$ identical particles with mass $m$ moving in a common potential $V(\mathbf{x})$. The possible energies of a single particle will then be given by the eigenvalues of the ordinary one-particle Hamiltonian

$$
\begin{equation*}
\widehat{h}=\frac{\widehat{\mathbf{p}}^{2}}{2 m}+U(\widehat{\mathbf{x}}) \tag{10.1}
\end{equation*}
$$

Using the one-particle identity operator $\hat{1}=\int d^{3} x|\mathbf{x}\rangle\langle\mathbf{x}|$, it can be written is the equivalent way $\widehat{h}=\int d^{3} x|\mathbf{x}\rangle h(\mathbf{x})\langle\mathbf{x}|$ where

$$
\begin{equation*}
h(\mathbf{x})=-\frac{\hbar^{2}}{2 m} \boldsymbol{\nabla}^{2}+U(\mathbf{x}) \tag{10.2}
\end{equation*}
$$

is the ordinary differential Hamiltonian operator. The corresponding eigenvalue equation $\widehat{h}|k\rangle=\varepsilon_{k}|k\rangle$ has orthogonal eigenstates with the standard normalization $\left\langle k \mid k^{\prime}\right\rangle=\delta_{k k^{\prime}}$. They also form a complete set, i.e.

$$
\begin{equation*}
\sum_{k=0}^{\infty}|k\rangle\langle k|=\widehat{1} \tag{10.3}
\end{equation*}
$$

In coordinate space the equivalent eigenvalue problem is given by $h(\mathbf{x}) u_{k}(\mathbf{x})=\varepsilon_{k} u_{k}(\mathbf{x})$ and corresponds to the orbital equation (9.16) in the previous chapter. But here we find it convenient to denote the corresponding eigenfunction by $u_{k}(\mathbf{x})=\langle\mathbf{x} \mid k\rangle$ when the particle is in the orbital or mode with quantum number $k$. From the above they are thus orthonormalized as

$$
\begin{equation*}
\int d^{3} x u_{k}^{*}(\mathbf{x}) u_{k^{\prime}}(\mathbf{x})=\delta_{k k^{\prime}} \tag{10.4}
\end{equation*}
$$

The completeness relation (10.3) similarly becomes

$$
\begin{equation*}
\sum_{k=0}^{\infty} u_{k}(\mathbf{x}) u_{k}^{*}\left(\mathbf{x}^{\prime}\right)=\delta\left(\mathbf{x}-\mathbf{x}^{\prime}\right) \tag{10.5}
\end{equation*}
$$

Needless to say, this is all valid for any dimension $d$ of the space we consider, but we have chosen to work in ordinary case with $d=3$.

The total energy $E$ of all the independent particles will now follow from the full Hamiltonian

$$
\begin{equation*}
\widehat{H}_{N}=\sum_{i=1}^{N} \widehat{h}_{i}=\widehat{h}_{1}+\widehat{h}_{2}+\cdots+\widehat{h}_{N} \tag{10.6}
\end{equation*}
$$

It is obviously invariant under all permutation of the particle labels, i.e. under the full permutation group. The energy of the system will now be the eigenvalues of the corresponding multiparticle Schrödinger equation $\widehat{H}_{N}\left|\Psi_{N}\right\rangle=E\left|\Psi_{N}\right\rangle$.
A possible eigenstate is obviously the direct product $\left|k_{1}\right\rangle\left|k_{2}\right\rangle \cdots\left|k_{N}\right\rangle$ where the first particle is in the state $\left|k_{1}\right\rangle$, the second particle in state $\left|k_{2}\right\rangle$ and so on. But it is not an eigenstate of the permutation group. For particles in three dimensions the physical states must be completely symmetric for bosons and completely antisymmetric for fermions. Let us denote these as $\left|k_{1} k_{2} \cdots k_{N}\right\rangle$. By combining all the $N$ ! permutations of the $N$ particle labels with the appropriate signs, we can write these as

$$
\begin{equation*}
\left|k_{1} k_{2} \cdots k_{N}\right\rangle=\sqrt{\frac{1}{N!}} \sum_{P} \epsilon^{P}\left|k_{P(1)}\right\rangle\left|k_{P(2)}\right\rangle \cdots\left|k_{P(N)}\right\rangle \tag{10.7}
\end{equation*}
$$

with $\epsilon=+1$ for bosons and $\epsilon=-1$ for fermions and where the sum goes over all particle permuations $P(i)$. With only $N=2$ particles these states are the usual symmetric and antisymmetric combinations we have met before. For three particles we can write the state as

$$
\begin{align*}
\left|k_{1} k_{2} k_{3}\right\rangle & =\sqrt{\frac{1}{6}}\left(\left|k_{1}\right\rangle\left|k_{2}\right\rangle\left|k_{3}\right\rangle+\epsilon\left|k_{1}\right\rangle\left|k_{3}\right\rangle\left|k_{2}\right\rangle+\left|k_{2}\right\rangle\left|k_{3}\right\rangle\left|k_{1}\right\rangle\right. \\
& \left.+\epsilon\left|k_{2}\right\rangle\left|k_{1}\right\rangle\left|k_{3}\right\rangle+\left|k_{3}\right\rangle\left|k_{1}\right\rangle\left|k_{2}\right\rangle+\epsilon\left|k_{3}\right\rangle\left|k_{2}\right\rangle\left|k_{1}\right\rangle\right) \tag{10.8}
\end{align*}
$$

For fermions the three quantum numbers $k_{1}, k_{2}$ and $k_{3}$ must all be different. This just reflects the Pauli principle saying that there can be no more than one particle in each quantum state. In this case the symmetrized state is also seen to be normalized to one. With bosons there can be an arbitrary number of particles in each state and two or all three of the quantum numbers $k_{1}, k_{2}$ and $k_{3}$ can be equal. The state $\left|k_{1} k_{2} k_{3}\right\rangle$ is
then no longer normalized to one. This is most easily seen when $k_{1}=k_{2}=k_{3} \equiv k$ since $|k k k\rangle=\sqrt{6}|k\rangle|k\rangle|k\rangle$. This enhancement of the bosonic norm can be said to be a reflection of the phenomenon of Bose condensation.
A convenient basis in this $N$-particle Hilbert space is the coordinate basis $\left|\mathbf{x}_{1} \mathbf{x}_{2} \cdots \mathbf{x}_{N}\right\rangle$ formed by the similarly symmetrized products of $\left|\mathbf{x}_{1}\right\rangle\left|\mathbf{x}_{2}\right\rangle \cdots\left|\mathbf{x}_{N}\right\rangle$. They satisfy the completeness relation

$$
\begin{equation*}
\frac{1}{N!} \int d x_{1} \int d x_{2} \cdots \int d x_{N}\left|\mathbf{x}_{1} \mathbf{x}_{2} \cdots \mathbf{x}_{N}\right\rangle\left\langle\mathbf{x}_{N} \cdots \mathbf{x}_{2} \mathbf{x}_{1}\right|=1 \tag{10.9}
\end{equation*}
$$

where we must divide by $N$ ! because of the permutation invariance of the integrand. For the system in a general state $|\Psi\rangle$ we can then define the corresponding wavefunction by the matrix element $\left\langle\mathbf{x}_{1} \mathbf{x}_{2} \cdots \mathbf{x}_{N} \mid \Psi\right\rangle$. In particular, for the state (10.7) we have the wavefunction $\Psi_{N}\left(\mathbf{x}_{1}, \mathbf{x}_{2}, \ldots, \mathbf{x}_{N}\right)=\left\langle\mathbf{x}_{1} \mathbf{x}_{2} \cdots \mathbf{x}_{N} \mid k_{1} k_{2} \cdots k_{N}\right\rangle$. Writing this out in detail, one finds that the result can be written on the compact form

$$
\Psi_{N}\left(\mathbf{x}_{1}, \mathbf{x}_{2}, \cdots, \mathbf{x}_{N}\right)=\left|\begin{array}{cccc}
u_{k_{1}}\left(\mathbf{x}_{1}\right) & u_{k_{2}}\left(\mathbf{x}_{1}\right) & \cdots & u_{k_{N}}\left(\mathbf{x}_{1}\right)  \tag{10.10}\\
\vdots & & & \vdots \\
u_{k_{1}}\left(\mathbf{x}_{N}\right) & u_{k_{2}}\left(\mathbf{x}_{N}\right) & \cdots & u_{k_{N}}\left(\mathbf{x}_{N}\right)
\end{array}\right|_{\epsilon}
$$

using the basic matrix element $u_{k_{n}}\left(\mathbf{x}_{i}\right)=\left\langle\mathbf{x}_{i} \mid k_{n}\right\rangle$. The object on the right-hand side is a determinant in the case of fermions with $\epsilon=-1$. Comparing with (9.17) in the previous chapter, we see that it is just the Slater determinant. For bosons it is called a permanent and is evaluated exactly as a determinant except that all the different products of its elements are combined with only positive signs.

When we bring the $n$th particle to the first position in the vector $\left|k_{1} k_{2} \cdots k_{N}\right\rangle$ in (10.7)we must exchange it with $n-1$ other particles. We then pick up a phase depending on the state being bosonic or fermionic. More specifically, we have

$$
\begin{equation*}
\left|k_{1} k_{2} \cdots k_{n} \cdots k_{N}\right\rangle=\epsilon^{n-1}\left|k_{n} k_{1} k_{2} \cdots \underline{k_{n}} \cdots k_{N}\right\rangle \tag{10.11}
\end{equation*}
$$

where the underlined particle is absent. For fermions this corresponds to exchanging rows in the Slater determinant.

The energy of such a physical state is then $E=\varepsilon_{k_{1}}+\varepsilon_{k_{2}}+\cdots+\varepsilon_{k_{N}}$. When the number of particles is very large, it is obviously very cumbersome to write out the exact state of each particle. This is in particular true for the physically most important multiparticle states which are those with lowest energies. In the bosonic case most of the particles will be in the one-particle ground state $u_{0}$ and thus have the same quantum numbers. From our knowledge of the harmonic oscillator and the quantized electromagnetic field we know what to do. We introduce also here the occupation numbers $n_{k}$ which give the number of particles in the one-particle mode $u_{k}$. The multiparticle state

$$
\begin{equation*}
\left|\Psi_{N}\right\rangle=\left|n_{0}, n_{1}, n_{2}, \cdots, n_{k}, \cdots\right\rangle \tag{10.12}
\end{equation*}
$$

has then $n_{0}$ particles in the lowest one-particle state $\left|k_{0}\right\rangle, n_{1}$ particles in the first excited state $\left|k_{1}\right\rangle$ and so on. For bosons these occupation numbers can take any integer value $n_{k}=(0,1,2, \ldots)$ while for fermions we can only have $n_{k}=(0,1)$. A couple of such states are shown in the figure. The total number of particles in the state (10.12) is then

$$
\begin{equation*}
N=\sum_{k=0}^{\infty} n_{k}=n_{0}+n_{1}+n_{2}+\cdots \tag{10.13}
\end{equation*}
$$


a)

b)

Figure 10.1: Occupation number representations of the bosonic state $|4,3,2,1,2,0,1\rangle$ in a) and the fermionic state $|1,1,0,1,1,0,1\rangle$ in b$)$.
while the total energy becomes

$$
\begin{equation*}
E=\sum_{k=0}^{\infty} n_{k} \varepsilon_{k}=n_{0} \varepsilon_{0}+\varepsilon_{1} n_{1}+\varepsilon_{2} n_{2}+\cdots \tag{10.14}
\end{equation*}
$$

Now we only have to give a few occupation numbers in order to give a complete quantum specification of the multiparticle system when it is in a low-energy state. What remains to do, is the construction of these states with the correct symmetry properties.

### 10.2 Fock space

The symmetrized states $\left|k_{1} k_{2} \cdots k_{N}\right\rangle$ are vectors in the Hilbert space $\mathcal{H}_{N}$ of $N$ identical particles. It is orthognal to the corresponding Hilbert spaces of systems with different number of particles. Each of these suffices as long as the number of particles remain fixed. But in many important and interesting physical systems in Nature this is not so as we know from emission of absorbtion of photons. In order to study these more general phenomena we must join all the separate Hilbert spaces for fixed particle numbers into a more general Hilbert space accommodating systems with an arbitrary number of particles. It is called a Fock space and is formally given by the infinite, direct product

$$
\begin{equation*}
\mathcal{F}=\mathcal{H}_{0} \otimes \mathcal{H}_{1} \otimes \mathcal{H}_{2} \otimes \cdots \otimes \mathcal{H}_{N} \otimes \cdots \tag{10.15}
\end{equation*}
$$

We will then be doing quantum field theory instead of manybody quantum mechanics. Even for systems with a small and fixed number of particles we will find the fieldtheoretic formulation very elegant and powerful.
Formally we can join the Hilbert spaces with fixed particle numbers by introducing new operators that connect them. For this purpose we define a operator $\hat{a}_{k}^{\dagger}$ which creates
the $N+1$ particle state $\left|k k_{1} k_{2} \cdots k_{N}\right\rangle$ from the $N$ particle state $\left|k_{1} k_{2} \cdots k_{N}\right\rangle$,

$$
\begin{equation*}
\widehat{a}_{k}^{\dagger}\left|k_{1} k_{2} \cdots k_{N}\right\rangle=\left|k k_{1} k_{2} \cdots k_{N}\right\rangle \tag{10.16}
\end{equation*}
$$

The new particle with quantum number $k$ is seen to be added in the position of the first particle. Acting on the empty state $|0\rangle$ it creates the one-particle state

$$
\begin{equation*}
\widehat{a}_{k}^{\dagger}|0\rangle=|k\rangle \tag{10.17}
\end{equation*}
$$

For each one-particle state there is a corresponding creation operator. The application of two creation operators gives similarly

$$
\hat{a}_{k}^{\dagger} \widehat{a}_{k^{\prime}}^{\dagger}\left|k_{1} k_{2} \cdots k_{N}\right\rangle=\left|k k^{\prime} k_{1} \cdots k_{N}\right\rangle
$$

But from the definition of the symmetrized states we have $\left|k k^{\prime} k_{1} \cdots k_{N}\right\rangle=\epsilon\left|k^{\prime} k k_{1} \cdots k_{N}\right\rangle$. Thus it follows that $\widehat{a}_{k}^{\dagger} \widehat{a}_{k^{\prime}}^{\dagger}=\epsilon \widehat{a}_{k^{\prime}}^{\dagger} \widehat{a}_{k}^{\dagger}$. This result can be written in the more standard form as

$$
\begin{equation*}
\left[\widehat{a}_{k}^{\dagger}, \widehat{a}_{k^{\prime}}^{\dagger}\right]_{\epsilon}=0 \tag{10.18}
\end{equation*}
$$

where the generalized commutator is defined as $[\widehat{A}, \widehat{B}]_{\epsilon}=\widehat{A} \widehat{B}-\epsilon \widehat{B} \widehat{A}$. For bosons with $\epsilon=+1$ this is just the ordinary commutator, while for fermions it is the anticommutator. Usually it is written with curly brackets, i.e.

$$
\begin{equation*}
\{\widehat{A}, \widehat{B}\}=\widehat{A} \widehat{B}+\widehat{B} \widehat{A} \tag{10.19}
\end{equation*}
$$

It was first encountered in the mathematics of the Pauli spin matrices as seen from (5.80).

By repeated applications of creation operators on the empty state, we can create any many-particle state which should then automatically have the right symmetry properties. Taking the adjoint of (10.18), we get the corresponding general commutator between annihilation operators,

$$
\begin{equation*}
\left[\widehat{a}_{k}, \widehat{a}_{k^{\prime}}\right]_{\epsilon}=0 \tag{10.20}
\end{equation*}
$$

which also reflects directly the symmetry properties of the states.
The adjoint operator $\widehat{a}_{k}$ correspondingly destroys a particle in the $N$ particle state $\left|k_{1} k_{2} \cdots k_{N}\right\rangle$ and takes it into a state with $N-1$ particles. When it acts on the empty state, there is nothing to annihilate and it gives zero,

$$
\begin{equation*}
\widehat{a}_{k}|0\rangle=0 \tag{10.21}
\end{equation*}
$$

Since $\langle k|=\langle 0| \widehat{a}_{k}$, we must then also have

$$
\begin{equation*}
\widehat{a}_{k}\left|k^{\prime}\right\rangle=\delta_{k k^{\prime}}|0\rangle \tag{10.22}
\end{equation*}
$$

since $\left\langle k \mid k^{\prime}\right\rangle=\delta_{k k^{\prime}}$. Usually the empty state has lowest energy without real particles. This is the quantum the vacuum state.
Writing the above result as $\widehat{a}_{k}\left|k^{\prime}\right\rangle=\left\langle k \mid k^{\prime}\right\rangle|0\rangle$, we can now find the action of the annihilation operator on a general $N$-particle state. It will act on each particle in the
state, removing one after the other. But the particle to be annihilated must be brought to the front position in the state vector. Using then the basic symmetry requirement (10.11), we get

$$
\begin{equation*}
\widehat{a}_{k}\left|k_{1} k_{2} \cdots k_{N}\right\rangle=\sum_{n=1}^{N} \epsilon^{n-1}\left\langle k \mid k_{n}\right\rangle\left|k_{1} k_{2} \cdots \underline{k_{n}} \cdots k_{N}\right\rangle \tag{10.23}
\end{equation*}
$$

The annihilation operator is thus seen to act like a derivation operator, picking up a $\operatorname{sign} \epsilon$ as it works its way through the many-particle state.

In order to find the commutator between creation and annihilation operators, we first apply an annihilation operator to (10.16) and then use (10.23) to obtain

$$
\begin{aligned}
\widehat{a}_{k} \hat{a}_{k^{\prime}}^{\dagger}\left|k_{1} k_{2} \cdots k_{N}\right\rangle & =\widehat{a}_{k}\left|k^{\prime} k_{1} k_{2} \cdots k_{N}\right\rangle \\
& =\left\langle k \mid k^{\prime}\right\rangle\left|k_{1} k_{2} \cdots k_{N}\right\rangle+\sum_{n=1}^{N} \epsilon^{n}\left\langle k \mid k_{n}\right\rangle\left|k^{\prime} k_{1} \cdots \underline{k_{n}} \cdots k_{N}\right\rangle
\end{aligned}
$$

When the product of operators is applied in the opposite order, it similarly follows that

$$
\begin{aligned}
\hat{a}_{k^{\prime}}^{\dagger} \widehat{a}_{k}\left|k_{1} k_{2} \cdots k_{N}\right\rangle & =\widehat{a}_{k^{\prime}}^{\dagger} \sum_{n=1}^{N} \epsilon^{n-1}\left\langle k \mid k_{n}\right\rangle\left|k_{1} \cdots \underline{k_{n}} \cdots k_{N}\right\rangle \\
& =\sum_{n=1}^{N} \epsilon^{n-1}\left\langle k \mid k_{n}\right\rangle\left|k^{\prime} k_{1} \cdots \underline{k_{n}} \cdots k_{N}\right\rangle
\end{aligned}
$$

Multiplying this last expression by $\epsilon$ and comparing with the result above, we find immediately the general commutator result

$$
\begin{equation*}
\left[\widehat{a}_{k}, \widehat{a}_{k^{\prime}}^{\dagger}\right]_{\epsilon}=\left\langle k \mid k^{\prime}\right\rangle=\delta_{k k^{\prime}} \tag{10.24}
\end{equation*}
$$

For bosons we see that it agrees with the corresponding commutators for the quanta in the harmonic oscillator and for photons. In fact, this canonical commutator lies at the heart of all quantum field theories, non-relativistic as here but also in relativistic theories like quantum electrodynamics. We thus conclude that bosons are quantized with commutators and fermions with anticommutators. This is the crucial, mathematical difference between the two families of particles.

The normalized multiparticle state (10.12) is now constructed as

$$
\begin{equation*}
\left|\Psi_{N}\right\rangle=\frac{\left(\widehat{a}_{0}^{\dagger}\right)^{n_{0}}}{\sqrt{n_{0}!}} \frac{\left(\widehat{a}^{\dagger}\right)^{n_{1}}}{\sqrt{n_{1}!}} \frac{\left(\widehat{a}_{2}^{\dagger}\right)^{n_{2}}}{\sqrt{n_{2}!}} \cdots \frac{\left(\widehat{a}_{k}^{\dagger}\right)^{n_{k}}}{\sqrt{n_{k}!}} \cdots|0\rangle \tag{10.25}
\end{equation*}
$$

as for photons. For fermions the occupations numbers would just take the two values $n_{k}=(0,1)$. They are eigenvalues of the number operators $\widehat{n}_{k}=\widehat{a}_{k}^{\dagger} \widehat{a}_{k}$. Indeed, with help of our basic commutators, we verify that

$$
\begin{equation*}
\left[\widehat{n}_{k}, \widehat{a}_{k^{\prime}}^{\dagger}\right]=\widehat{a}_{k}^{\dagger} \delta_{k k^{\prime}}, \quad\left[\widehat{n}_{k}, \widehat{a}_{k^{\prime}}\right]=-\widehat{a}_{k} \delta_{k k^{\prime}} \tag{10.26}
\end{equation*}
$$

both for bosons and fermions, using the identity $[\widehat{A} \widehat{B}, \widehat{C}]=\widehat{A}[\widehat{B}, \widehat{C}]_{\epsilon}+\epsilon[\widehat{A}, \widehat{C}]_{\epsilon} \widehat{B}$. Thus we have number operators with the wanted effect

$$
\begin{equation*}
\widehat{n}_{k}\left|n_{0}, n_{1}, n_{2}, \cdots, n_{k}, \cdots\right\rangle=n_{k}\left|n_{0}, n_{1}, n_{2}, \cdots, n_{k}, \cdots\right\rangle \tag{10.27}
\end{equation*}
$$

The total number operator $\widehat{N}$ with the eigenvalue (10.13) is therefore

$$
\begin{equation*}
\widehat{N}=\sum_{k=0}^{\infty} \widehat{n}_{k}=\sum_{k=0}^{\infty} \widehat{a}_{k}^{\dagger} \widehat{a}_{k} \tag{10.28}
\end{equation*}
$$

It can be used both for closed systems where the total particle number is fixed and for open systems where this number changes. The latter situation is found in statistical physics for systems in thermal and chemical equlibrium with the environment.

Similarly, the Hamiltonian for the full system of particles is

$$
\begin{equation*}
\widehat{H}=\sum_{k=0}^{\infty} \varepsilon_{k} \widehat{a}_{k}^{\dagger} \widehat{a}_{k} \tag{10.29}
\end{equation*}
$$

since it must reproduce the energy eigenvalue (10.14). It gives the time evolution of the system in the Fock space described by the state vector $|\Psi, t\rangle$ as dictated by the Schrödinger equation

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t}|\Psi, t\rangle=\widehat{H}|\Psi, t\rangle \tag{10.30}
\end{equation*}
$$

valid for all quantum systems.

### 10.3 Quantum field operators

The one-particle states $|k\rangle$ have formed the basis from which we can struct all multiparticle states in the full Fock space. Creation and annihilation operators for other choices of basis states can now be constructed by standard transformation theory. A special and very important rôle is played by the operators which create and annihilate particles in position eigenstates $|\mathbf{x}\rangle$. These are easily found from the completeness relation (10.3) which gives

$$
\begin{equation*}
|\mathbf{x}\rangle=\sum_{k=0}^{\infty} u_{k}^{*}(\mathbf{x})|k\rangle \tag{10.31}
\end{equation*}
$$

Since we now can write $|k\rangle=\widehat{a}_{k}^{\dagger}|0\rangle$, we therefore have $|\mathbf{x}\rangle=\widehat{\psi}^{\dagger}(\mathbf{x})|0\rangle$ where

$$
\begin{equation*}
\widehat{\psi}^{\dagger}(\mathbf{x})=\sum_{k=0}^{\infty} u_{k}^{*}(\mathbf{x}) \widehat{a}_{k}^{\dagger} \tag{10.32}
\end{equation*}
$$

is a new operator. It could just as well have been called $\widehat{a}_{\mathbf{x}}^{\dagger}$ since it is seen to create a particle in the position $\mathbf{x}$. The corresponding Hermitean conjugate operator

$$
\begin{equation*}
\widehat{\psi}(\mathbf{x})=\sum_{k=0}^{\infty} \widehat{a}_{k} u_{k}(\mathbf{x}) \tag{10.33}
\end{equation*}
$$

similarly annihilates a particle in the same position. Needless to say, this is the quantized field operator for the non-relativistic particles we consider and it acts in Fock space. All the particles are identical since they are described by the same field.

We can now express the previous operator equations in terms if these new field operators. For instance, the canonical commutator (10.24) is replaced by

$$
\begin{aligned}
{\left[\widehat{\psi}(\mathbf{x}), \widehat{\psi}^{\dagger}\left(\mathbf{x}^{\prime}\right)\right]_{\epsilon} } & =\sum_{k k^{\prime}} u_{k}(\mathbf{x}) u_{k^{\prime}}^{*}\left(\mathbf{x}^{\prime}\right)\left[\widehat{a}_{k}, \widehat{a}_{k^{\prime}}^{\dagger}\right]_{\epsilon}=\sum_{k k^{\prime}} u_{k}(\mathbf{x}) u_{k^{\prime}}^{*}\left(\mathbf{x}^{\prime}\right) \delta_{k k^{\prime}} \\
& =\sum_{k=0}^{\infty} u_{k}(\mathbf{x}) u_{k}^{*}\left(\mathbf{x}^{\prime}\right)
\end{aligned}
$$

From the completeness relation (10.5) we thus have

$$
\begin{equation*}
\left[\widehat{\psi}(\mathbf{x}), \widehat{\psi}^{\dagger}\left(\mathbf{x}^{\prime}\right)\right]_{\epsilon}=\delta\left(\mathbf{x}-\mathbf{x}^{\prime}\right) \tag{10.34}
\end{equation*}
$$

By a similar calculation we find the commutator between two field operators be zero, i.e. $\left[\widehat{\psi}(\mathbf{x}), \widehat{\psi}\left(\mathbf{x}^{\prime}\right)\right]_{\epsilon}=0$. It is a direct reflection of the zero commutator between any two annihilation operators $\widehat{a}_{k}$ and $\widehat{a}_{k^{\prime}}$.
Similarly, we can also express the previous operators $\widehat{N}$ and $\widehat{H}$ in terms of these new field operators. Since they both involve the products of creation and annihilation operators, consider the integral

$$
\int d^{3} x \widehat{\psi}^{\dagger}(\mathbf{x}) \widehat{\psi}(\mathbf{x})=\sum_{k k^{\prime}} \widehat{a}_{k}^{\dagger} \widehat{a}_{k^{\prime}} \int d^{3} x u_{k}^{*}(\mathbf{x}) u_{k^{\prime}}(\mathbf{x})
$$

Using here the former normalization integral (10.4), the right-hand side just simplifies to $(10.28)$ for the number operator. Thus we have

$$
\begin{equation*}
\widehat{N}=\int d^{3} x \widehat{\psi}^{\dagger}(\mathbf{x}) \widehat{\psi}(\mathbf{x}) \tag{10.35}
\end{equation*}
$$

It is therefore appropriate to call the local operator $\widehat{\rho}(\mathbf{x})=\widehat{\psi}^{\dagger}(\mathbf{x}) \widehat{\psi}(\mathbf{x})$ for the number density operator.

In exactly the same way we can write the Hamiltonian (10.29) as

$$
\begin{equation*}
\widehat{H}=\int d^{3} x \widehat{\psi}^{\dagger}(\mathbf{x})\left[-\frac{1}{2 m} \nabla^{2}+U(\mathbf{x})\right] \widehat{\psi}(\mathbf{x}) \tag{10.36}
\end{equation*}
$$

since the action of the differential operator $h(\mathbf{x})$ within the brackets picks out the eigenvalue $\varepsilon_{k^{\prime}}$ when it acts on the mode-function $u_{k^{\prime}}(\mathbf{x})$ in $\widehat{\psi}(\mathbf{x})$. Here it is therefore also appropriate to define a Hamiltonian density operator $\widehat{\mathcal{H}}(\mathbf{x})=\widehat{\psi}^{\dagger}(\mathbf{x}) h(\mathbf{x}) \widehat{\psi}(\mathbf{x})$. It determines the energy density of particles at point $\mathbf{x}$.
The $N$-particle state $\widehat{\psi}^{\dagger}\left(\mathbf{x}_{1}\right) \widehat{\psi}^{\dagger}\left(\mathbf{x}_{2}\right) \cdots \widehat{\psi}^{\dagger}\left(\mathbf{x}_{N}\right)|0\rangle$ now decribes particles with definite positions in coordinate space and with the right symmetry properties. For different values of $N$ these form a basis in Fock space similar to the coordinate basis $\left|\mathbf{x}_{1} \mathbf{x}_{2} \cdots \mathbf{x}_{N}\right\rangle$ we initially used. It is therefore natural to define the many-particle wavefunction for the state (10.25) as the matrix element

$$
\begin{equation*}
\Psi_{N}\left(\mathbf{x}_{1}, \mathbf{x}_{2}, \ldots, \mathbf{x}_{N}\right)=\langle 0| \widehat{\psi}\left(\mathbf{x}_{1}\right) \widehat{\psi}\left(\mathbf{x}_{2}\right) \ldots \widehat{\psi}\left(\mathbf{x}_{N}\right)\left|\Psi_{N}\right\rangle \tag{10.37}
\end{equation*}
$$

When $N=1$ this is almost obvious and can easily be shown for $N=2$, reproducing the correct symmetric combination of one-particle wavefunctions for bosons and antisymmetric for fermions. In the general case we reproduce the Slater determinant-permanent in (10.10).

From the Schrödinger equation (10.30) we can now find the time dependence of these field operators in the usual way. Since the Hamiltonian is constant in time, we have

$$
\begin{equation*}
\widehat{\psi}(\mathbf{x}, t)=e^{i \widehat{H} t / \hbar} \widehat{\psi}(\mathbf{x}) e^{-i \widehat{H} t / \hbar} \tag{10.38}
\end{equation*}
$$

Using Lie's formula (2.53) and the commutators in (10.26), this can be evaluated to give

$$
\begin{align*}
\widehat{\psi}(\mathbf{x}, t) & =\sum_{k=0}^{\infty} u_{k}(\mathbf{x})\left[\widehat{a}_{k}+\left[\widehat{H}, \widehat{a}_{k}\right](i t / \hbar)+\frac{1}{2!}\left[\widehat{H},\left[\widehat{H}, \widehat{a}_{k}\right]\right](i t / \hbar)^{2}+\ldots\right] \\
& =\sum_{k=0}^{\infty} \widehat{a}_{k} u_{k}(\mathbf{x}) e^{-i \varepsilon_{k} t / \hbar} \tag{10.39}
\end{align*}
$$

Notice that from the definition (10.38), the canonical commutator (10.34) is also valid for the time-dependent field $\widehat{\psi}(\mathbf{x}, t)$ as long as the two operators in the commutator are taken at the same time. It is then said to be an equal time commutator.

Calculating in the same way the time evolution of the Hamiltonian (10.36), we get on the left side the expected result $\widehat{H}(t)=\widehat{H}$ since the operator $\widehat{H}$ commutes with itself. This should be expected because we consider a isolated system where the total energy is constant. However, on the right-hand side we obtain $\widehat{H}=\int d^{3} x \widehat{\psi}^{\dagger}(\mathbf{x}, t) h(\mathbf{x}) \widehat{\psi}(\mathbf{x}, t)$ which can be useful in some instances. Here $h(\mathbf{x})$ is the differential one-particle Hamiltonian operator (10.2) used previously.

With this time dependence we are really in the Heisenberg picture as explained in Chapter 2. An expression equivalent to the time-dependence in (10.39) can therefore be derived directly from the Heisenberg equation (2.48). For the field operator it gives

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t} \widehat{\psi}(\mathbf{x}, t)=[\widehat{\psi}(\mathbf{x}, t), \widehat{H}] \tag{10.40}
\end{equation*}
$$

With the above expression for the Hamiltonian, it then follows that

$$
\begin{aligned}
i \hbar \frac{\partial}{\partial t} \widehat{\psi}(\mathbf{x}, t) & =\int d^{3} x^{\prime}\left[\widehat{\psi}(\mathbf{x}, t), \widehat{\psi}^{\dagger}\left(\mathbf{x}^{\prime}, t\right) h\left(\mathbf{x}^{\prime}\right) \widehat{\psi}\left(\mathbf{x}^{\prime}, t\right)\right] \\
& \left.=\int d^{3} x^{\prime} \delta\left(\mathbf{x}-\mathbf{x}^{\prime}\right) h\left(\mathbf{x}^{\prime}\right) \widehat{\psi}\left(\mathbf{x}^{\prime}, t\right)\right]
\end{aligned}
$$

where the non-zero contribution comes from the canonical commutator (10.34). We thus obtain the equation of motion

$$
\begin{equation*}
i \hbar \frac{\partial}{\partial t} \widehat{\psi}(\mathbf{x}, t)=\left[-\frac{\hbar^{2}}{2 m} \boldsymbol{\nabla}^{2}+U(\mathbf{x})\right] \widehat{\psi}(\mathbf{x}, t) \tag{10.41}
\end{equation*}
$$

for the field operator. It is seen to satisfy the same Schrödinger equation as the corresponding one-particle wavefunction $\psi(\mathbf{x}, t)$ does. In some way the wavefunction have become an operator. For this reason non-relativistic quantum field theory is sometimes called second quantization.

When the motion of a single particle is first quantized, we find the allowed values of energy, momentum etc of each particle. But with many identical particles present in the same system, there is a new layer of quantum conditions being enforced which makes the
complex-valued wavefuncton into an operator in Fock space. The many-particle system is therefore said to be second quantized when described in this field-theoretic language.

But one should never confuse wavefunctions with quantum fields. The former can be functions of several coordinate vectors describing the positions of many particles while a field is always a function of just one coordinate vector. A more modern name for the formalism is Schrödinger quantum field theory since the field operator satisfies the Schrödinger equation (10.41).
For non-interacting particles the Hamiltonian (10.1) simplifies to just $\widehat{h}_{0}=\widehat{\mathbf{p}}^{2} / 2 m$. The one-particle solutions are then just the free states $|\mathbf{k}\rangle$ characterized by the wavenumber $\mathbf{k}$ and with eigenvalues $\varepsilon_{\mathbf{k}}=(\hbar \mathbf{k})^{2} / 2 m$. In the coordinate basis these are the plane waves

$$
\begin{equation*}
u_{\mathbf{k}}(\mathbf{x})=\langle\mathbf{x} \mid \mathbf{k}\rangle=\sqrt{\frac{1}{V}} e^{i \mathbf{k} \cdot \mathbf{x}} \tag{10.42}
\end{equation*}
$$

when the system is in a finite box $V$. These were also used for the quantization of the electromagnetic field in Chapter 7 where the normalization integral (7.64) was derived correponding to (10.4) in this chapter.

Everything from the previous section can obviously be taken directly over to this special case. The canonical commutator is $\left[\widehat{a}_{\mathbf{k}}, \widehat{a}_{\mathbf{k}^{\prime}}^{\dagger}\right]_{\epsilon}=\delta_{\mathbf{k k}^{\prime}}$ while the field Hamiltonian $\widehat{H}_{0}=\sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} \widehat{a}_{\mathbf{k}}^{\dagger} \widehat{a}_{\mathbf{k}}$ has the same form as the corresponding one-particle Hamiltonian $\widehat{h}_{0}=\sum_{\mathbf{k}} \varepsilon_{\mathbf{k}}|\mathbf{k}\rangle\langle\mathbf{k}|$. For the field operator (10.39) we now get

$$
\begin{equation*}
\widehat{\psi}(\mathbf{x}, t)=\sqrt{\frac{1}{V}} \sum_{\mathbf{k}} \widehat{a}_{\mathbf{k}} e^{i\left(\mathbf{k} \cdot \mathbf{x}-\varepsilon_{\mathbf{k}} t / \hbar\right)} \tag{10.43}
\end{equation*}
$$

and only valid in this case of free particles. This corresponds to the first part of the electromagnetic field operator (7.86). The second part is missing here since $\widehat{\psi}(\mathbf{x}, t)$ is a quantum field operator for non-relativstic particles. Photons are always relativistic.

We can make use of this plane-wave expansion also in the more general case with interactions. This situation would arise if we could not solve the more general one-body mode equation $h(\mathbf{x}) u_{k}(\mathbf{x})=\varepsilon_{k} u_{k}(\mathbf{x})$. But these exact modes are really not absolutely necessary for the derivation of the field operator $\widehat{\psi}(\mathbf{x})$. In the definition (10.31) we only made use of the fact that they form a complete set. Thus we can always choose to use instead the plane-wave $u_{\mathbf{k}}$, thus giving

$$
\begin{equation*}
\widehat{\psi}(\mathbf{x})=\sqrt{\frac{1}{V}} \sum_{\mathbf{k}} \widehat{a}_{\mathbf{k}} e^{i \mathbf{k} \cdot \mathbf{x}} \tag{10.44}
\end{equation*}
$$

for the field operator (10.33). However, the Hamiltonian which now has the more general form (10.36), becomes

$$
\begin{equation*}
\widehat{H}=\sum_{\mathbf{k}}\left[\varepsilon_{\mathbf{k}} \widehat{a}_{\mathbf{k}}^{\dagger} \widehat{a}_{\mathbf{k}}+\sum_{\mathbf{q}} \widehat{a}_{\mathbf{k}+\mathbf{q}}^{\dagger} \widehat{a}_{\mathbf{k}} U_{\mathbf{q}}\right] \tag{10.45}
\end{equation*}
$$

with the expansion (10.44). Here $U_{\mathbf{q}}$ is the Fourier transform

$$
U_{\mathbf{q}}=\frac{1}{V} \int d^{3} x e^{-i \mathbf{q} \cdot \mathbf{x}} U(\mathbf{x})
$$

We can understand the last term in the Hamiltonian as a scattering process where an incoming particle with momentum $\hbar \mathbf{k}$ hits the potential $U(\mathbf{x})$, receives a momentum $\hbar \mathbf{q}$ from it and continues as a new particle with momentum $\hbar \mathbf{k}^{\prime}=\hbar \mathbf{k}+\hbar \mathbf{q}$, leaving the potential. This is depicted in the figure which is called a Feynman diagram for


Figure 10.2: Feynman diagram for particle interacting with an external potential.
this scattering process. As for photons, we see that in this second-quantized picture the particle at an interaction point is first annihilated and then instantanously recreated as a new particle at the same point. With first quantization in ordinary quantum mechanics we would say that it is the same particle that moves throughout the scattering process.

For free particles we could calculate exactly the general time dependence (10.43) of the field operator. This we can no longer do for an interacting field. That should be clear from the derivation in (10.39) where the commutator $\left[\widehat{H}, \widehat{a}_{k}\right]$ now will an infinite number of terms, making it impossible to find a closed expression for the sum in the Lie formula.

This reflects the fact that we no longer have the exact solution of the interacting oneparticle problem. We must therefore include the effect of the external potential $U(\mathbf{x})$ in some other way. However, when it is sufficiently weak, we can use standard perturbation theory. It then becomes clear that the perturbative vacuum state $|0\rangle$ satisfying $\widehat{a}_{\mathbf{k}}|0\rangle=0$ no longer is the physical vaccum state $|\Omega\rangle$, i.e. the one with lowest energy. This new ground state will in general be a Fock state of the generic form (10.25) or a superposition of such states. In general we will then obviously also have $\widehat{a}_{\mathbf{k}}|\Omega\rangle \neq 0$. For the same reason, in the plane-wave expansion (10.44) of the field operator, the interacting N particle wavefunction will now be $\langle\Omega| \widehat{\psi}\left(\mathbf{x}_{1}\right) \widehat{\psi}\left(\mathbf{x}_{2}\right) \ldots \widehat{\psi}\left(\mathbf{x}_{N}\right)\left|\Psi_{N}\right\rangle$ instead of (10.37). In the generic case it is a complicated, infinite sum or integral over products of plane waves.

### 10.4 Inter-particle interactions

So far we have only considered independent particles, i.e. without mutual interactions. Even if they interact with an external potential, they can considered to be free. And for this reason such a system is not very realistic and therefore just a starting point, particularly suitable for building up the necessary formalism.

It really important physics in real systems found around us is due to the forces between the particles. In most cases these are two-body interactions described by a potential $V\left(\mathbf{x}-\mathbf{x}^{\prime}\right)$ depending only on the separation between the particles. In quantum mechanics the corresponding two-body operator is then

$$
\begin{equation*}
\widehat{v}=\frac{1}{2} \int d^{3} x \int d^{3} x^{\prime}\left|\mathbf{x} \mathbf{x}^{\prime}\right\rangle V\left(\mathbf{x}-\mathbf{x}^{\prime}\right)\left\langle\mathbf{x}^{\prime} \mathbf{x}\right| \tag{10.46}
\end{equation*}
$$

when one particle is at position $\mathbf{x}$ and the other at $\mathbf{x}^{\prime}$. The factor $1 / 2$ in front is a normalization constant which follows from the completeness relation (10.9) for the special case $V=1$.

Let us now construct the corresponding interaction operator in the field-theoretic formulation. Writing $\left|\mathbf{x x}^{\prime}\right\rangle=\widehat{\psi}^{\dagger}(\mathbf{x}) \widehat{\psi}^{\dagger}\left(\mathbf{x}^{\prime}\right)|0\rangle$ in (10.46), we must have

$$
\begin{equation*}
\widehat{V}=\frac{1}{2} \int d^{3} x \int d^{3} x^{\prime} \widehat{\psi}^{\dagger}(\mathbf{x}) \widehat{\psi}^{\dagger}\left(\mathbf{x}^{\prime}\right) V\left(\mathbf{x}-\mathbf{x}^{\prime}\right) \widehat{\psi}\left(\mathbf{x}^{\prime}\right) \widehat{\psi}(\mathbf{x}) \tag{10.47}
\end{equation*}
$$

when we follow the construction we previously used for one-body operator like the Hamiltonian. A Feynman diagram describing this interaction is shown in Fig.10.3. On the left side at position $\mathbf{x}$ a particle is annihilated by the operator $\widehat{\psi}(\mathbf{x})$ and then


Figure 10.3: Feynman diagram for a two-particle interaction via a potential.
recreated by $\widehat{\psi}^{\dagger}(\mathbf{x})$. This happens simultaneously also at the other vertex $\mathbf{x}^{\prime}$ of the interaction. Expressing the field operators in terms of annihilation and creation operators as given in (10.33), we obtain

$$
\begin{equation*}
\widehat{V}=\frac{1}{2} \sum_{k l m n} V_{k l ; m n} \widehat{a}_{k}^{\dagger} \widehat{a}_{l}^{\dagger} \widehat{a}_{m} \widehat{a}_{n} \tag{10.48}
\end{equation*}
$$

where the interaction matrix element

$$
\begin{equation*}
V_{k l ; m n}=\int d^{3} x \int d^{3} x^{\prime} u_{k}^{*}(\mathbf{x}) u_{l}^{*}\left(\mathbf{x}^{\prime}\right) V\left(\mathbf{x}-\mathbf{x}^{\prime}\right) u_{m}\left(\mathbf{x}^{\prime}\right) u_{n}(\mathbf{x}) \tag{10.49}
\end{equation*}
$$

The ordering of the operators is important. As a check of this formula, consider the special case that the potential is constant $V=1$. Then we have $V_{k l ; m n}=\delta_{k n} \delta_{l m}$ so that

$$
\widehat{V}=\frac{1}{2} \sum_{k l} \widehat{a}_{k}^{\dagger} \widehat{a}_{l}^{\dagger} \widehat{a}_{l} \widehat{a}_{k}
$$

Here the total number operator $\widehat{N}$ appears. With the annihilation operator it has the commutator $\left[\widehat{N}, \widehat{a}_{k}\right]=-\widehat{a}_{k}$ and therefore

$$
\widehat{V}=\frac{1}{2} \sum_{k} \widehat{a}_{k}^{\dagger} \widehat{N} \widehat{a}_{k}=\frac{1}{2} \sum_{k} \widehat{a}_{k}^{\dagger}\left(\widehat{a}_{k} \widehat{N}-\widehat{a}_{k}\right)=\frac{1}{2}\left(\widehat{N}^{2}-\widehat{N}\right)
$$

This is correct since there are $N(N-1) / 2$ interacting pairs in the system.
Including this mutual interaction, the previous Hamiltonian (10.36) now becomes

$$
\begin{align*}
\widehat{H} & =\int d^{3} x \widehat{\psi}^{\dagger}(\mathbf{x})\left[-\frac{1}{2 m} \nabla^{2}+U(\mathbf{x})\right] \widehat{\psi}(\mathbf{x}) \\
& +\frac{1}{2} \int d^{3} x \int d^{3} x^{\prime} \hat{\psi}^{\dagger}(\mathbf{x}) \widehat{\psi}^{\dagger}\left(\mathbf{x}^{\prime}\right) V\left(\mathbf{x}-\mathbf{x}^{\prime}\right) \widehat{\psi}\left(\mathbf{x}^{\prime}\right) \widehat{\psi}(\mathbf{x}) \tag{10.50}
\end{align*}
$$

In terms of creation and annihilation operators it takes the more compact form

$$
\begin{equation*}
\widehat{H}=\sum_{k=0}^{\infty} \widehat{a}_{k}^{\dagger} \widehat{a}_{k} \varepsilon_{k}+\frac{1}{2} \sum_{k l m n} V_{k l ; m n} \widehat{a}_{k}^{\dagger} \widehat{a}_{l}^{\dagger} \widehat{a}_{m} \widehat{a}_{n} \tag{10.51}
\end{equation*}
$$

and has the general form used in condensed matter and low-energy nuclear physics. Indeed, using the variational principle one can with this Hamiltonian derive more systematically the approximate Hartree-Fock equations for the mode functions written down at the end of the previous chapter. Notice again that the physical vacuum state $|\Omega\rangle$ of this interacting system is in general different from the perturbative vacuum $|0\rangle$ defined by $\widehat{a}_{k}|0\rangle=0$.

As already mentioned, one has a many-particle wavefunction $\Psi_{N}\left(\mathbf{x}_{1}, \mathbf{x}_{2}, \ldots, \mathbf{x}_{N}\right)$ also with interacting particles using the definition (10.37) with the perturbative vacuum $|0\rangle$ replaced by the physical vacuum or ground state $|\Omega\rangle$. It is then a reassuring problem to show that it satisfies the ordinary, $N$-particle Schrödinger wave equation with the interparticle potential $V\left(\mathbf{x}-\mathbf{x}^{\prime}\right)$. For instance, for $N=2$ and discarding the external potential $U$, one finds

$$
\begin{equation*}
\left[-\frac{\hbar^{2}}{2 m}\left(\boldsymbol{\nabla}_{1}^{2}+\boldsymbol{\nabla}_{2}^{2}\right)+V\left(\mathbf{x}_{1}-\mathbf{x}_{2}\right)\right] \Psi_{2}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right)=E \Psi_{2}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right) \tag{10.52}
\end{equation*}
$$

when the energy of the physical vacuum is set to zero, i.e. $\widehat{H}|\Omega\rangle=0$.
Many important systems consist of particles interacting over a finite range. As was explained in Chapter 6 in connection with scattering of such particles, the detailed shape of the potential is then not important at sufficiently low energies. Only the scattering length $a$ has an observable consequence. When this is the case, we can replace the above physical potential with a mathematical potential which is simpler to calculate with. And the smartest choice is to use the Dirac $\delta$-function, i.e.

$$
\begin{equation*}
V\left(\mathbf{x}-\mathbf{x}^{\prime}\right) \rightarrow g \delta\left(\mathbf{x}-\mathbf{x}^{\prime}\right) \tag{10.53}
\end{equation*}
$$

Here $g$ is an effective coupling constant which can be expressed in terms of the scattering length $a$. In fact, from (??) and its definition, we find $g=2 \pi a \hbar^{2} / \mu$ where $\mu=m / 2$ is the reduced mass of the two interacting particles.

With this replacement of the potential in the Hamiltonian (10.50), it becomes a volume integral over the Hamiltonian density

$$
\begin{equation*}
\widehat{\mathcal{H}}=\widehat{\psi}^{\dagger}\left(-\frac{\hbar^{2}}{2 m} \boldsymbol{\nabla}^{2}+U\right) \widehat{\psi}+\frac{1}{2} g \widehat{\psi}^{\dagger} \widehat{\psi}^{\dagger} \widehat{\psi} \widehat{\psi} \tag{10.54}
\end{equation*}
$$

Calculating now the equation of motion of $\widehat{\psi}(\mathbf{x})$ from (10.40) as before, the result is obviously no longer the ordinary Schrödinger equation, but instead

$$
\begin{equation*}
i \hbar \frac{\partial \widehat{\psi}}{\partial t}=\left[-\frac{\hbar^{2}}{2 m} \nabla^{2}+U+g \widehat{\psi}^{\dagger} \widehat{\psi}\right] \widehat{\psi} \tag{10.55}
\end{equation*}
$$

Since it involves an extra, cubic term, it is sometimes called the non-linear Schrödinger equation. But what we have here is a dynamical equation for the field while the Schrödinger equation (10.30) for the state vector is always linear.

### 10.5 Fermions

In order to distinguish these from bosons, it is here convenient to change notation to $\widehat{a}_{k} \rightarrow \widehat{b}_{k}$ for the fermion operators. The mode expansion (10.33) is then

$$
\begin{equation*}
\widehat{\psi}(\mathbf{x})=\sum_{k=0}^{\infty} \widehat{b}_{k} u_{k}(\mathbf{x}) \tag{10.56}
\end{equation*}
$$

These annihilation operators now satisfy $\left\{\widehat{b}_{k}, \widehat{b}_{k^{\prime}}\right\}=0$. In particular we have $\widehat{b}_{k}^{2}=0$ when acting on any state. This is again a direct reflection of the Pauli principle saying that there can be no more than one particle in each state.
Squaring the number operator $\widehat{n}_{k}=\hat{b}_{k}^{\dagger} \widehat{b}_{k}$ one obtains $\widehat{n}_{k}^{2}=\widehat{b}_{k}^{\dagger} \widehat{b}_{k} \hat{b}_{k}^{\dagger} \widehat{b}_{k}$. Interchanging the two middle operators give $\widehat{n}_{k}^{2}=\widehat{b}_{k}^{\dagger}\left(1-\widehat{b}_{k}^{\dagger} \widehat{b}_{k}\right) \widehat{b}_{k}$ using the canonical anticommutator $\left\{\widehat{b}_{k}, \hat{b}_{k^{\prime}}^{\dagger}\right\}=\delta_{k k^{\prime}}$. The last term is now seen to be zero. The number operator thus satisfies $\widehat{n}_{k}\left(\widehat{n}_{k}-1\right)=0$ and thus has only the eigenvalues $n_{k}=(0,1)$ again as Pauli said.

So far we have just considered spinless fermions. But the spin-statistics theorem says that fermions must have half-integer spins. Electrons have spin $S=1 / 2$ and must therefore be equipped with an extra quantum number $\sigma=\uparrow$ or $\sigma=\downarrow$ depending on the spin direction. In a plane-wave basis we will then need the both quantum number $\mathbf{k}$ and $\sigma$ so that $\widehat{b}_{k}=\widehat{b}_{\mathbf{k} \sigma}$. The canonical anticommutator is then

$$
\begin{equation*}
\left\{\widehat{b}_{\mathbf{k} \sigma}, \widehat{b}_{\mathbf{k}^{\prime} \sigma^{\prime}}^{\dagger}\right\}=\delta_{\mathbf{k} \mathbf{k}^{\prime}} \delta_{\sigma \sigma^{\prime}} \tag{10.57}
\end{equation*}
$$

while $\left\{\widehat{b}_{\mathbf{k} \sigma}, \widehat{b}_{\mathbf{k}^{\prime} \sigma^{\prime}}\right\}=0$. The field operator (10.56) becomes a Pauli spinor

$$
\begin{equation*}
\widehat{\psi}(\mathbf{x})=\binom{\widehat{\psi}_{\uparrow}(\mathbf{x})}{\widehat{\psi}_{\downarrow}(\mathbf{x})} \tag{10.58}
\end{equation*}
$$

where the components have the expansion

$$
\begin{equation*}
\widehat{\psi}_{\sigma}(\mathbf{x})=\sqrt{\frac{1}{V}} \sum_{\mathbf{k}} \widehat{b}_{\mathbf{k} \sigma} e^{i \mathbf{k} \cdot \mathbf{x}} \tag{10.59}
\end{equation*}
$$

in analogy with (10.44) for bosons. From (10.34) we then have the canonical anticommutator in coordinate space

$$
\begin{equation*}
\left\{\widehat{\psi}_{\sigma}(\mathbf{x}, t), \widehat{\psi}_{\sigma^{\prime}}^{\dagger}\left(\mathbf{x}^{\prime}, t\right)\right\}=\delta\left(\mathbf{x}-\mathbf{x}^{\prime}\right) \delta_{\sigma \sigma^{\prime}} \tag{10.60}
\end{equation*}
$$

for the non-relativistic spinor field.
The one-particle spin operator $\mathbf{S}=(\hbar / 2) \boldsymbol{\sigma}=(\hbar / 2) \int d^{3} x|\mathbf{x}\rangle \boldsymbol{\sigma}\langle\mathbf{x}|$ will now become the second quantized operator

$$
\begin{equation*}
\widehat{\mathbf{S}}=\frac{1}{2} \hbar \int d^{3} x \widehat{\psi}^{\dagger}(\mathbf{x}) \boldsymbol{\sigma} \widehat{\psi}(\mathbf{x}) \tag{10.61}
\end{equation*}
$$

using the same recipe as previously. For instance, the $z$-component is given by the field operator

$$
\begin{equation*}
\widehat{S}_{z}=\frac{1}{2} \hbar \int d^{3} x \widehat{\psi}_{\sigma}^{\dagger}(\mathbf{x})\left(\sigma_{z}\right)_{\sigma \sigma^{\prime}} \widehat{\psi}_{\sigma^{\prime}}(\mathbf{x})=\frac{1}{2} \hbar \sum_{\mathbf{k}}\left(\widehat{b}_{\mathbf{k} \uparrow}^{\dagger} \widehat{b}_{\mathbf{k} \uparrow}-\widehat{b}_{\mathbf{k} \downarrow}^{\dagger} \widehat{b}_{\mathbf{k} \downarrow}\right) \tag{10.62}
\end{equation*}
$$

which follows after the volume integration. The first term is seen to give $+\hbar / 2$ for each fermion with spin up while the last gives $-\hbar / 2$ from those with spin down. Similarly, the spin raising operator is $\widehat{S}_{+}=\hbar \sum_{\mathbf{k}} \widehat{b}_{\mathbf{k} \uparrow}^{\dagger} \widehat{b}_{\mathbf{k} \downarrow}$. When it acts on a one-particle state $|\mathbf{k} \downarrow\rangle=\widehat{b}_{\mathbf{k} \downarrow}^{\dagger}|0\rangle$ with spin down, it thus changes its spin to up, i.e. $\widehat{S}_{+}|\mathbf{k} \downarrow\rangle=|\mathbf{k} \uparrow\rangle$.

In an external potential the Hamiltonian of the fermions is still given by (10.36). The analogous Hamiltonian in an external electromagnetic field we can similarly be read off directly from the Pauli Hamiltonian (7.30) in Chapter 7.

For a realistic system we must also include inter-particle interactions as already described by (10.48) when we ignore the external potential. In a plane-wave basis we then evaluate the matrix element (10.49) and get the Hamiltonian

$$
\begin{equation*}
\widehat{H}=\sum_{\mathbf{k} \sigma} \varepsilon_{\mathbf{k}} \widehat{b}_{\mathbf{k} \sigma}^{\dagger} \widehat{b}_{\mathbf{k} \sigma}+\frac{1}{2} \sum_{\mathbf{k} \mathbf{q k} \mathbf{k}^{\prime}} V_{\mathbf{q}} \hat{b}_{\mathbf{k}+\mathbf{q}}^{\dagger} \widehat{b}_{\mathbf{k}^{\prime}-\mathbf{q}}^{\dagger} \widehat{b}_{\mathbf{k}^{\prime}} \widehat{b}_{\mathbf{k}} \tag{10.63}
\end{equation*}
$$

where $\varepsilon_{\mathbf{k}}=k^{2} / 2 m$ again are the one-particle energies. In the last term we have ignored the spin labels since these will be determined by the spin dependence of the potential. As illustrated by the Feynman digram in Fig.10.4, the interaction describes two particles with momenta $\mathbf{k}$ and $\mathbf{k}^{\prime}$ which interact via the Fourier component $V_{\mathbf{q}}$ of the potensial. After the interaction they have momenta $\mathbf{k}+\mathbf{q}$ and $\mathbf{k}^{\prime}-\mathbf{q}$ respectively.

Superconductivity was explained by Bardeen, Cooper and Schrieffer by a simpler version of this general Hamiltonian. It was based on the conjecture by Cooper that under special conditions in a metal two electrons will feel a weak, attractive force when they have opposite momenta and spin. This is called a Cooper pair. The corresponding interaction potential can be taken to be constant, i.e. $V_{\mathbf{q}} \rightarrow-V$. It can then be written as

$$
\begin{equation*}
\widehat{H}_{B C S}=\sum_{\mathbf{k}} \varepsilon_{\mathbf{k}}\left[\widehat{b}_{\mathbf{k}}^{\dagger} \widehat{b}_{\mathbf{k}}+\widehat{b}_{-\mathbf{k}}^{\dagger} \widehat{b}_{-\mathbf{k}}\right]-V \sum_{\mathbf{k} \mathbf{k}^{\prime}} \widehat{b}_{\mathbf{k}^{\prime}}^{\dagger} \widehat{b}_{-\mathbf{k}^{\prime}}^{\dagger} \widehat{b}_{\mathbf{k}} \widehat{b}_{-\mathbf{k}} \tag{10.64}
\end{equation*}
$$

when each wavenumber $\mathbf{k}$ is associated with spin up $\uparrow$ and $-\mathbf{k}$ implies spin down $\downarrow$. On this form one can find the physical ground state $|\Omega\rangle$ which has lower energy than the perturbative vacuum $|0\rangle$. It can be considered to be a condensate of Cooper pairs that conducts electricity with zero resistance.


Figure 10.4: Feynman diagram for a two-particle interaction in momentum space.

### 10.6 Particles and holes

In the ground state of free fermions all the lowest one-particles states are occupied up to maximum momentum $\hbar k_{F}$ which is called the Fermi momentum. The corresponding energy $\varepsilon_{F}=\hbar^{2} k_{F}^{2} / 2 m$ is the Fermi energy. It is a spherical surface in momentum space. For a real metal with electrons moving in an ionic lattice, it will have a more complicated geometrical shape.

The ground state $|F\rangle$ can now be constructed as

$$
\begin{equation*}
|F\rangle=\prod_{|\mathbf{k}| \leq k_{F}} \hat{b}_{\mathbf{k}}^{\dagger}|0\rangle \tag{10.65}
\end{equation*}
$$

when we here and in the following ignore the spin labels. It has an energy

$$
\begin{equation*}
E=\langle F| \widehat{H}|F\rangle=\sum_{|\mathbf{k}| \leq k_{F}} \varepsilon_{\mathbf{k}} \tag{10.66}
\end{equation*}
$$

This is just one example of similar sums of the general form

$$
\begin{equation*}
S=\sum_{\mathbf{k}} f\left(\varepsilon_{\mathbf{k}}\right) \tag{10.67}
\end{equation*}
$$

which occurs very often in statistical and solid state physics. Defining the density of states by

$$
\begin{equation*}
D(\varepsilon)=\frac{1}{V} \sum_{\mathbf{k}} \delta\left(\varepsilon-\varepsilon_{\mathbf{k}}\right) \tag{10.68}
\end{equation*}
$$

where $V$ is the volume of the system, the sum follows then from

$$
\begin{equation*}
S=V \int d \varepsilon D(\varepsilon) f(\varepsilon) \tag{10.69}
\end{equation*}
$$

This integral is usually simpler to do than the original when one knows the density of states.

For free fermions $\varepsilon_{\mathbf{k}}=\hbar^{2} k^{2} / 2 m$ it is easily found to be

$$
\begin{align*}
D(\varepsilon) & =\int \frac{d^{3} k}{(2 \pi)^{3}} \delta\left(\varepsilon-\hbar^{2} k^{2} / 2 m\right)=\frac{m}{\pi^{2}} \int_{0}^{\infty} d k k^{2} \delta((\hbar k-\sqrt{2 m \varepsilon})(\hbar k+\sqrt{2 m \varepsilon})) \\
& =\frac{m}{2 \pi^{2} \hbar^{3}} \sqrt{2 m \varepsilon} \tag{10.70}
\end{align*}
$$

where the contribution to the integral comes from the zero $\hbar k=\sqrt{2 m \varepsilon}$ in the deltafunction. In two dimensions we would similarly have obtained $D(\varepsilon)=m / 2 \pi \hbar^{2}$.
The ground state energy (10.66) now follows as

$$
\begin{equation*}
E=V \int_{0}^{\varepsilon_{F}} d \varepsilon \varepsilon D(\varepsilon)=\frac{m V}{5 \pi^{2} \hbar^{3}} \sqrt{2 m} \varepsilon_{F}^{5 / 2} \tag{10.71}
\end{equation*}
$$

Similarly, the density $\rho=N / V$ of fermions is

$$
\begin{equation*}
\rho=\int_{0}^{\varepsilon_{F}} d \varepsilon D(\varepsilon)=\frac{m}{3 \pi^{2} \hbar^{3}} \sqrt{2 m} \varepsilon_{F}^{3 / 2} \tag{10.72}
\end{equation*}
$$

Combining these two results, we find for the average fermion energy $E / N=(3 / 5) \varepsilon_{F}$. In two dimensions it is $\varepsilon_{F} / 2$ instead.

When a particle is excited from below the Fermi surface to a vacant state above it, a hole is created among the states where the fermion used to be. The creation operator for such a hole with momentum $\mathbf{k}$ can therefore be defined to be

$$
\begin{equation*}
\widehat{d}_{\mathbf{k}}^{\dagger}=\widehat{b}_{-\mathbf{k}}, \quad|\mathbf{k}| \leq k_{F} \tag{10.73}
\end{equation*}
$$

We can then divide the set of fermion creation and annihilation operators into two, hole operators $\widehat{d}_{\mathbf{k}}$ and $\widehat{d}_{\mathbf{k}}^{\dagger}$ with $|\mathbf{k}| \leq k_{F}$ and particle operators $\widehat{b}_{\mathbf{k}}$ and $\widehat{b}_{\mathbf{k}}^{\dagger}$ with $|\mathbf{k}|>k_{F}$. From the fundamental anticommutator (10.57) we then obviously have

$$
\begin{aligned}
\left\{\hat{b}_{\mathbf{k}}, \hat{b}_{\mathbf{k}^{\prime}}^{\dagger}\right\} & =\delta_{\mathbf{k}, \mathbf{k}^{\prime}} & & |\mathbf{k}|,\left|\mathbf{k}^{\prime}\right|>k_{F} \\
\left\{\widehat{d}_{\mathbf{k}}, \widehat{d}_{\mathbf{k}^{\prime}}^{\dagger}\right\} & =\delta_{\mathbf{k}, \mathbf{k}^{\prime}} & & |\mathbf{k}|,\left|\mathbf{k}^{\prime}\right| \leq k_{F}
\end{aligned}
$$

and $\left\{\widehat{b}_{\mathbf{k}}, \hat{d}_{\mathbf{k}^{\prime}}^{\dagger}\right\}=0$ since their momenta will always be different.
In the mode sum for the field operator (10.59) we can now make a split into two physically distinct pieces,

$$
\begin{equation*}
\widehat{\psi}(\mathbf{x})=\sqrt{\frac{1}{V}}\left[\sum_{|\mathbf{k}|>k_{F}} \widehat{b}_{\mathbf{k}} e^{i \mathbf{k} \cdot \mathbf{x}}+\sum_{|\mathbf{k}| \leq k_{F}} \widehat{d}_{\mathbf{k}}^{\dagger} e^{-i \mathbf{k} \cdot \mathbf{x}}\right] \tag{10.74}
\end{equation*}
$$

where the first corresponds to particles and the last to holes. In this latter one we have made the replacement $\mathbf{k} \rightarrow-\mathbf{k}$ in the summation.
For the free Hamiltonian $\widehat{H}=\sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} \widehat{b}_{\mathbf{k}}^{\dagger} \widehat{b}_{\mathbf{k}}$ we we can make a similar separation between the two contributions,

$$
\begin{align*}
\widehat{H} & =\sum_{|\mathbf{k}|>k_{F}} \varepsilon_{\mathbf{k}} \widehat{b}_{\mathbf{k}}^{\dagger} \widehat{b}_{\mathbf{k}}+\sum_{|\mathbf{k}| \leq k_{F}} \varepsilon_{\mathbf{k}} \widehat{b}_{\mathbf{k}}^{\dagger} \widehat{b}_{\mathbf{k}} \\
& =E_{F}+\sum_{|\mathbf{k}|>k_{F}} \varepsilon_{\mathbf{k}} \widehat{b}_{\mathbf{k}}^{\dagger} \widehat{b}_{\mathbf{k}}-\sum_{|\mathbf{k}| \leq k_{F}} \varepsilon_{\mathbf{k}} \widehat{d}_{\mathbf{k}}^{\dagger} \widehat{d}_{\mathbf{k}} \tag{10.75}
\end{align*}
$$

after using the anti-commutator between the two hole operators. We thus see that the energies of the hole excitations are negative when measured with respect to the Fermi energy. In a relativistic theory the existence of antiparticles come about in a similar way. Indeed, an antiparticle is really the absence of a particle with positive energy. But one should be careful with this analogy since antiparticles exist both for bosons and fermions.

### 10.7 Lattice field theory

So far we have ignored the background potential $U(\mathbf{x})$ in which the electrons are moving. In solid state physics this is formed by the positively charged ions situated on a regular lattice. It has translational invariance as already discussed in Chapter 3. As was demonstrated there, this symmetry has important implications for the energy eigenstates of the particles. One finds energy bands separated by forbidden gaps.

When the coupling to the ions become strong, the electrons are most likely to be found near the corresponding lattice sites where the ions are situated. They can move by


Figure 10.5: Two-dimensional square lattice where each site has four nearest neighbours.
jumping to one of the nearby sites. Usually these are taken to be the nearest neighbouring sites. For the quadratic lattice in two dimensions shown in Fig. 10.5 there are four such sites, each separated from the original site $\mathbf{x}$ by the lattice constant $a$. Their positions are therefore $\mathbf{x}^{\prime}=\mathbf{x}+a \mathbf{e}_{x}, \mathbf{x}-a \mathbf{e}_{x}, \mathbf{x}+a \mathbf{e}_{y}, \mathbf{x}-a \mathbf{e}_{y}$ where $\mathbf{e}_{i}$ are unit vectors along the crystal axes. For a cubic lattice in three dimensions there are six such nearest neighbours.

A particle on such a three-dimensional lattice has then a one-particle Hamiltonian (3.46) which we now write as

$$
\begin{equation*}
\widehat{h}=E_{0} \sum_{\mathbf{x}}|\mathbf{x}\rangle\langle\mathbf{x}|-A \sum_{\left\langle\mathbf{x}, \mathbf{x}^{\prime}\right\rangle}|\mathbf{x}\rangle\left\langle\mathbf{x}^{\prime}\right| \tag{10.76}
\end{equation*}
$$

where $E_{0}$ is the energy of the particle localized at a site. In the last term $-A$ is the hopping amplitude and the sum goes over all sites $\mathbf{x}$ and their nearest neighbours $\mathbf{x}^{\prime}$.

We can now use the same construction as in the previous sections to find the corresponding second quantized Hamiltonian describing an arbitrary number of free particles on the lattice. If we let the operator $\widehat{a}_{\mathbf{x}}^{\dagger}$ create a particle at site $\mathbf{x}$ so that we can write $|\mathbf{x}\rangle=\widehat{a}_{\mathbf{x}}^{\dagger}|0\rangle$, the many-particle Hamiltonian will be

$$
\begin{equation*}
\widehat{H}=E_{0} \sum_{\mathbf{x}} \widehat{a}_{\mathbf{x}}^{\dagger} \widehat{a}_{\mathbf{x}}-A \sum_{<\mathbf{x}, \mathbf{x}^{\prime}>} \widehat{a}_{\mathbf{x}}^{\dagger} \widehat{a}_{\mathbf{x}^{\prime}} \tag{10.77}
\end{equation*}
$$

The canonical commutator is now

$$
\begin{equation*}
\left[\widehat{a}_{\mathbf{x}}, \widehat{a}_{\mathbf{x}^{\prime}}^{\dagger}\right]_{\epsilon}=\delta_{\mathbf{x x} \mathbf{x}^{\prime}} \tag{10.78}
\end{equation*}
$$

It is the discrete version of (10.34) valid in the continuous case. The operator $\widehat{n}_{\mathbf{x}}=\widehat{a}_{\mathbf{x}}^{\dagger} \widehat{a}_{\mathbf{x}}$ counts the number of particles at lattice site $\mathbf{x}$.

Because of the jumping term in the full Hamiltonian (10.78) the eigenstates will be superpositions involving all the localized states $|\mathbf{x}\rangle$. In fact, from (3.44) in Chapter 3 we know that an eigenstate with wavenumber $\mathbf{k}$ will be

$$
\begin{equation*}
|\mathbf{k}\rangle=\sum_{\mathbf{x}} e^{i \mathbf{k} \cdot \mathbf{x}}|\mathbf{x}\rangle=\sum_{\mathbf{x}} e^{i \mathbf{k} \cdot \mathbf{x}} \widehat{a}_{\mathbf{x}}^{\dagger}|0\rangle \tag{10.79}
\end{equation*}
$$

As a check, let us consider

$$
\widehat{H}|\mathbf{k}\rangle=E_{0} \sum_{\mathbf{x} \mathbf{x}^{\prime}} \widehat{a}_{\mathbf{x}}^{\dagger} \widehat{a}_{\mathbf{x}} \widehat{a}_{\mathbf{x}^{\prime}}^{\dagger} e^{i \mathbf{k} \cdot \mathbf{x}^{\prime}}|0\rangle-A \sum_{<\mathbf{x}, \mathbf{x}^{\prime}>} \widehat{a}_{\mathbf{x}}^{\dagger} \widehat{a}_{\mathbf{x}^{\prime}} \sum_{\mathbf{x}^{\prime \prime}} \widehat{a}_{\mathbf{x}^{\prime \prime}}^{\dagger} e^{i \mathbf{k} \cdot \mathbf{x}^{\prime \prime}}|0\rangle
$$

In the first term we interchange the operators $\widehat{a}_{\mathbf{x}}$ and $\widehat{a}_{\mathbf{x}^{\prime}}^{\dagger}$ using the commutator (10.78). This leaves $\delta_{\mathbf{x x}^{\prime}}$ since $\widehat{a}_{\mathbf{x}}|0\rangle=0$. Doing the same in the second term with the operators $\widehat{a}_{\mathbf{x}^{\prime}}$ and $\widehat{a}_{\mathbf{x}^{\prime \prime}}^{\dagger}$, leaves $\delta_{\mathbf{x}^{\prime} \mathbf{x}^{\prime \prime}}$. Thus

$$
\begin{aligned}
\widehat{H}|\mathbf{k}\rangle & =E_{0} \sum_{\mathbf{x}} \widehat{a}_{\mathbf{x}}^{\dagger} e^{i \mathbf{k} \cdot \mathbf{x}}|0\rangle-A \sum_{\left\langle\mathbf{x}, \mathbf{x}^{\prime}\right\rangle} \widehat{a}_{\mathbf{x}}^{\dagger} e^{i \mathbf{k} \cdot \mathbf{x}^{\prime}}|0\rangle \\
& =E_{0}|\mathbf{k}\rangle-A \sum_{i=1}^{3}\left(e^{i a \mathbf{k} \cdot \mathbf{e}_{i}}+e^{-i a \mathbf{k} \cdot \mathbf{e}_{i}}\right) \sum_{\mathbf{x}} \widehat{a}_{\mathbf{x}}^{\dagger} e^{i \mathbf{k} \cdot \mathbf{x}}|0\rangle
\end{aligned}
$$

and we have a result on the form $\widehat{H}|\mathbf{k}\rangle=\mathcal{E}_{\mathbf{k}}|\mathbf{k}\rangle$. This coherent excitation has therefore the energy

$$
\begin{equation*}
\mathcal{E}_{\mathbf{k}}=E_{0}-2 A \sum_{i=1}^{3} \cos k_{i} a \tag{10.80}
\end{equation*}
$$

in agreement with the result (3.48) and holds both for bosons and fermions on the lattice.

These eigenstates (10.79) can be written as $|\mathbf{k}\rangle=\widehat{a}_{\mathbf{k}}^{\dagger}|0\rangle$ where the corresponding creation operator is seen to be

$$
\begin{equation*}
\widehat{a}_{\mathbf{k}}^{\dagger}=\sqrt{\frac{1}{N}} \sum_{\mathbf{x}} \widehat{a}_{\mathbf{x}}^{\dagger} e^{i \mathbf{k} \cdot \mathbf{x}} \tag{10.81}
\end{equation*}
$$

where $N=\sum_{\mathbf{x}} 1$ now is the number of sites in the lattice. It is needed if we want to have the standard commutator $\left[\widehat{a}_{\mathbf{k}}, \widehat{a}_{\mathbf{k}^{\prime}}^{\dagger}\right]_{\epsilon}=\delta_{\mathbf{k k}^{\prime}}$ come out this way, as is easily verified. With these new creation and annihilation the lattice Hamiltonian (10.78) is diagonalized in the sense that the can be written as

$$
\begin{equation*}
\widehat{H}=\sum_{\mathbf{k}} \mathcal{E}_{\mathbf{k}} \widehat{a}_{\mathbf{k}}^{\dagger} \widehat{a}_{\mathbf{k}} \tag{10.82}
\end{equation*}
$$

The eigenvalues are therefore simply given in terms of the fermionic occupation numbers as $E=\sum_{\mathbf{k}} \mathcal{E}_{\mathbf{k}} n_{\mathbf{k}}$ where the corresponding eigenstates are constructed as in the generic case (10.25).
Electrons are charged and repel each other with Coulomb potential $V_{\mathbf{x x}^{\prime}}$ that increases when their separation decreases. This inter-particle interaction can now be described by the discretized version of (10.47). Combining it with the free Hamiltonian above, we have

$$
\begin{equation*}
\widehat{H}=E_{0} \sum_{\mathbf{x}} \widehat{a}_{\mathbf{x}}^{\dagger} \widehat{a}_{\mathbf{x}}-A \sum_{<\mathbf{x}, \mathbf{x}^{\prime}>} \widehat{a}_{\mathbf{x}}^{\dagger} \widehat{a}_{\mathbf{x}^{\prime}}+\frac{1}{2} \sum_{\mathbf{x x ^ { \prime }}} \widehat{a}_{\mathbf{x}}^{\dagger} \widehat{a}_{\mathbf{x}^{\prime}}^{\dagger} V_{\mathbf{x x}^{\prime}} \widehat{a}_{\mathbf{x}^{\prime}} \widehat{a}_{\mathbf{x}} \tag{10.83}
\end{equation*}
$$

The first term is like a chemical potential, the second term represents the kinetic energy while the second term is the potential or interaction energy.
A simpler version of this Hamiltonian which still incorporates the essential physics, is to include only the Coulomb repulsion for particles being on the same lattice sites. Without spin, the particles must then be bosons. Denoting this repulsive energy by $U$ and writing $\widehat{a}_{\mathbf{x}}^{\dagger} \widehat{a}_{\mathbf{x}}^{\dagger} \widehat{a}_{\mathbf{x}} \widehat{a}_{\mathbf{x}}=\widehat{n}_{\mathbf{x}} \widehat{n}_{\mathbf{x}}+\widehat{n}_{\mathbf{x}}$ valid for bosons, we then get

$$
\begin{equation*}
\widehat{H}=-A \sum_{<\mathbf{x}, \mathbf{x}^{\prime}>} \widehat{a}_{\mathbf{x}}^{\dagger} \widehat{a}_{\mathbf{x}^{\prime}}+U \sum_{\mathbf{x}} \widehat{n}_{\mathbf{x}} \widehat{n}_{\mathbf{x}} \tag{10.84}
\end{equation*}
$$

We have here ignored a diagonal term proportional to $E_{0}+U$ corresponding to the first term in (10.83). It acts only as a reference point for measuring energies. Even with all these simplifications only approximate results can be derived from this Hamiltonian when it is used to model quantum lattice particles in condensed matter physics.

With electrons we can only have at most two particles on the same site and only when they have opposite spins. Introducing again this quantum number, the above Hamiltonian must then be written more precisely as

$$
\begin{equation*}
\widehat{H}=-A \sum_{\left\langle\mathbf{x}, \mathbf{x}^{\prime}\right\rangle}\left(\widehat{b}_{\mathbf{x} \uparrow}^{\dagger} \widehat{b}_{\mathbf{x}^{\prime} \uparrow}+\widehat{b}_{\mathbf{x} \downarrow}^{\dagger} \widehat{b}_{\mathbf{x}^{\prime} \downarrow}\right)+U \sum_{\mathbf{x}} \widehat{n}_{\mathbf{x} \uparrow} \widehat{n}_{\mathbf{x} \downarrow} \tag{10.85}
\end{equation*}
$$

This particular version is called the Hubbard model and has been intensively studied as a model for strongly correlated electrons in metals exhibiting high- $T_{c}$ superconductivity. However, its detailed properties are to a large extent still unknown in spite of its very simple form.

