Quantum Mechanics: Fundamental Principles and Applications

John F. Dawson

Department of Physics, University of New Hampshire, Durham, NH 03824

October 14, 2009, 9:08am EST

Contents

List of Figures

List of Tables

Preface

In this book, I have tried to bridge the gap between material learned in an undergraduate course in quantum mechanics and an advanced relativistic field theory course. The book is a compilation of notes for a first year graduate course in non-relativistic quantum mechanics which I taught at the University of New Hampshire for a number of years. These notes assume an undergraduate knowledge of wave equation based quantum mechanics, on the level of Griffiths[1] or Liboff [2], and undergraduate mathematical skills on the level of Boas [3]. This book places emphasis on learning new theoretical methods applied to old non-relativistic ideas, with a eye to what will be required in relativistic field theory and particle physics courses. The result provides an introduction to quantum mechanics which is, I believe, unique.

The book is divided into two sections: Fundamental Principles and Applications. The fundamental principles section starts out in the usual way by reviewing linear algebra, vector spaces, and notation in the first chapter, and then in the second chapter, we discuss canonical quantization of classical systems. In the next two chapters, Path integrals and in- and out-states are discussed. Next is a chapter on the density matrix and Green functions in quantum mechanics where we also discuss thermal density matrices and Green functions. This is followed by a chapter on identical particles and second quantized non-relativistic fields. Next the Galilean group is discussed in detail and wave equations for massless and massive non-relativistic particles explored. Finally, the last chapter of the fundamental principles section is devoted to supersymmetry in non-relativistic quantum mechanics.

In the application section, I start by discussing finite quantum systems: the motion of electrons on molecules and on linear and circular chains. This is followed by chapters one and two dimensional wave mechanics and the WKB approximation. Then I discuss spin systems, the harmonic oscillator, and electrons and phonons on linear lattices. Approximation methods are discussed next, with chapters on perturbative and variational approximations. This is followed by a chapters on exactly solvable potential problems in non-relativistic quantum mechanics, and a detailed chapter on angular momentum theory in quantum mechanics. In the next chapter, we discuss several problems concerning the interactions of non-relativistic electrons with a classical electromagnetic fields, including the Hydrogen atom, and lastly, we include a chapter on scattering theory.

There are appendices giving operator identities, binomial coefficients, fourier transforms, and sections reviewing classical physics, differential geometry, classical statistical mechanics, and Schwinger's angular momentum theory.

Much of the material for these notes come from the many excellent books on the subject, and in many cases, I have only rearranged them in my own way. I have tried to give references to original material when this was done. Of course, any misunderstandings are my own.

I would like to thank ...

John Dawson
September, 2007
Durham, NH

References

[1] D. J. Griffiths, *Introduction to Quantum Mechanics* (Pearson, Prentice Hall, Upper Saddle River, NJ, 2005), second edition.

[2] R. L. Liboff, *Introductory Quantum Mechanics* (Addison-Wesley, awp:adr, 1997), third edition.

[3] M. L. Boaz, *Mathematical Methods in the Physical Sciences* (John Wiley & Sons, New York, NY, 1983).

Part I

Fundamental Principles

Chapter 1

Basic principles of quantum theory

Physical systems are represented in quantum theory by a complex **vector space** $\mathcal{V}$ with an inner product. The state of the system is described by a particular **vector** $|\,\Psi\,\rangle$ in this space. All the possible states of the system are represented by **basis vectors** in this space. Observables are represented by **Hermitian operators** acting in this vector space. The possible values of these observables are the **eigenvalues** of these operators. Probability amplitudes for observing these values are **inner products**. Symmetries of the physical system are represented by **unitary transformations** of the basis vectors in $\mathcal{V}$. All this is summarized in Table 1.1 below. Thus it will be important for us to study linear vector spaces in detail, which is the subject of this chapter.

Nature	*Quantum Theory*
The physical system	a vector space $\mathcal{V}$
The state of the system	a vector $\|\,\Psi\,\rangle$ in $\mathcal{V}$
All possible states of the system	a set of basis vectors
Observables	Hermitian operators
Possible values of observables	eigenvalues of operators
Probability amplitudes for events	inner products
Symmetries	unitary transformations

Table 1.1: Relation between Nature and Quantum Theory

1.1 Linear vector spaces

In the following[1], we will denote **scalars** (complex numbers) by a, b, c, ..., and **vectors** by $|\,\alpha\,\rangle$, $|\,\beta\,\rangle$, $|\,\gamma\,\rangle$, ...

Definition 1 (linear vector space)**.** A linear vector space $\mathcal{V}$ is a set of objects called vectors ($|\,\alpha\,\rangle$, $|\,\beta\,\rangle$, $|\,\gamma\,\rangle$, ...) which are **closed** under addition and scalar multiplication. That is, if $|\,\alpha\,\rangle$ and $|\,\beta\,\rangle$ are in $\mathcal{V}$, then $a|\,\alpha\,\rangle + b|\,\beta\,\rangle$ is in $\mathcal{V}$.

Vectors addition and scalar multiplication have commutative, associative, and distributive properties:

1. $|\,\alpha\,\rangle + |\,\beta\,\rangle = |\,\beta\,\rangle + |\,\alpha\,\rangle$. *commutative law*

[1] Much of the material in this chapter was taken from Serot [1, chapter 1]

2. $(|\,\alpha\,\rangle + |\,\beta\,\rangle) + |\,\gamma\,\rangle = |\,\alpha\,\rangle + (|\,\beta\,\rangle + |\,\gamma\,\rangle)$. *associative law*
3. $a(b|\,\alpha\,\rangle) = (ab)|\,\alpha\,\rangle$. *associative law*
4. $(a+b)|\,\alpha\,\rangle = a|\,\alpha\,\rangle + b|\,\alpha\,\rangle$. *distributive law*
5. $a(|\,\alpha\,\rangle + |\,\beta\,\rangle) = a|\,\alpha\,\rangle + a|\,\beta\,\rangle$. *distributive law*
6. There is a unique vector $|\,0\,\rangle$ in $\mathcal{V}$, called the **null vector**, with the properties, $|\,\alpha\,\rangle + |\,0\,\rangle = |\,\alpha\,\rangle$ and $0|\,\alpha\,\rangle = |\,0\,\rangle$, for all $|\,\alpha\,\rangle$.

Example 1 $(\mathbb{C}^N)$**.** The set of N complex numbers $(c_1, c_2, \ldots, c_N)$, where $c_i \in \mathbb{C}$. Addition of vectors is defined by addition of the components, and scalar multiplication by the multiplication of each element by the scalar. We usually write vectors as column matrices:

$$|\,c\,\rangle = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} \tag{1.1}$$

Example 2 $(\mathcal{P}^N)$**.** The set of all real polynomials $c(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_N t^N$ of degree less than N in an independent real variable t, $-1 \le t \le 1$. A vector is defined by $|\,\mathbf{c}\,\rangle = c(t)$. Addition and multiplication by a scalar are the ordinary ones for polynomials. Note that in this example, we define a secondary variable $t \in \mathbb{R}$, which is not in the vector space.

Example 3 $(C[a,b])$**.** The set of all continuous complex functions of a real variable on the closed interval $[a,b]$. Thus $|\,f\,\rangle = f(x)$, $a \le x \le b$. Again, in this example, we have a secondary base variable consisting of a real variable $x \in \mathbb{R}$.

1.2 Linear independence

A set of vectors $|\,e_1\,\rangle, |\,e_2\,\rangle, \ldots, |\,e_N\,\rangle$, are **linearly independent** if the relation,

$$\sum_{n=1}^{N} c_n\,|\,e_n\,\rangle = 0\,, \tag{1.2}$$

can only be true if: $c_n = 0$, $n = 1, \ldots, N$. Otherwise, the set of vectors are linearly dependent, which means that one of them can be expressed as a linear combination of the others.

The maximum number N of linearly independent vectors in a vector space $\mathcal{V}$ is called the **dimension** of the space, in which case the set of vectors provides a **basis** set for $\mathcal{V}$. Any vector in the space can be written as a linear combination of the basis vectors. We can easily prove this:

Theorem 1. *Let $|\,e_n\,\rangle$, $n = 1, \ldots, N$, be a basis in $\mathcal{V}$. Then any vector $|\,\alpha\,\rangle$ in $\mathcal{V}$ can be represented by:*

$$|\,\alpha\,\rangle = \sum_{n=1}^{N} a_n|\,e_n\,\rangle\,,$$

where a_n are complex numbers.

Proof. Since $|\,\alpha\,\rangle$ and $|\,e_n\,\rangle$, $n = 1, \ldots, N$ are $N+1$ vectors in $\mathcal{V}$, they must be linearly *dependent*. So there must exist complex numbers c, c_n, $n = 1, \ldots, N$, not all zero, such that

$$c\,|\,\alpha\,\rangle + \sum_{n=1}^{N} c_n|\,e_n\,\rangle = 0\,.$$

But $c \neq 0$, otherwise the set $|\, e_n \,\rangle$, $n = 1, \ldots, N$, would be linearly dependent, which they are not. Therefore

$$|\, \alpha \,\rangle = \sum_{n=1}^{N} \frac{-c_n}{c} |\, e_n \,\rangle = \sum_{n=1}^{N} a_n |\, e_n \,\rangle \,.$$

where $a_n = -c_n/c$. If a different set of coefficients b_n, $n = 1, \ldots, N$ existed, we would then have by subtraction,

$$\sum_{n=1}^{N} (a_n - b_n) |\, e_n \,\rangle = 0 \,,$$

which can be true only if $b_n = a_n$, $n = 1, \ldots, N$, since the set $|\, e_n \,\rangle$ is linearly dependent. Thus the components a_n are *unique* for the basis set $|\, e_n \,\rangle$. □

In this chapter, we mostly consider linear vector spaces which have *finite* dimensions. Our examples 1 and 2 above have dimension N, whereas example 3 has infinite dimensions.

1.3 Inner product

An inner product maps pairs of vectors to complex numbers. It is written: $g(|\, \alpha \,\rangle, |\, \beta \,\rangle)$. That is, g is a function with two slots for vectors which map each pair of vectors in $\mathcal{V}$ to a complex number. The inner product must be defined so that it is **anti-linear** with respect to the first argument and **linear** with respect to the second argument. Because of this linearity and anti-linearity property, it is useful to write the inner product simply as:

$$g(|\, \alpha \,\rangle, |\, \beta \,\rangle) \equiv \langle\, \alpha \,|\, \beta \,\rangle \,. \tag{1.3}$$

The inner product must have the properties:

1. $\langle\, \alpha \,|\, \beta \,\rangle = \langle\, \beta \,|\, \alpha \,\rangle^* =$ a complex number.
2. $\langle\, a\,\alpha + b\,\beta \,|\, \gamma \,\rangle = a^* \langle\, \alpha \,|\, \gamma \,\rangle + b^* \langle\, \beta \,|\, \gamma \,\rangle$.
3. $\langle\, \gamma \,|\, a\,\alpha + b\,\beta \,\rangle = a \langle\, \gamma \,|\, \alpha \,\rangle + b \langle\, \gamma \,|\, \beta \,\rangle$,
4. $\langle\, \alpha \,|\, \alpha \,\rangle \geq 0$, with the equality holding only if $|\, \alpha \,\rangle = |\, 0 \,\rangle$.

The **norm**, or length, of a vector is defined by $\|\alpha\|^2 \equiv \langle\, \alpha \,|\, \alpha \,\rangle > 0$. A **Hilbert** space is a linear vector space with an inner product for each pair of vectors in the space.

Using our examples of linear vector spaces, one possible definition of the inner products is:

Example 4 ($\mathbb{C}^N$)**.** For example:

$$\langle\, a \,|\, b \,\rangle = a_1^* b_1 + a_2^* b_2 + \cdots + a_N^* b_N \,, \tag{1.4}$$

Example 5 ($\mathcal{P}^N$)**.** We can take:

$$\langle\, a \,|\, b \,\rangle = \int_{-1}^{+1} a^*(t)\, b(t)\, \mathrm{d}t \,, \tag{1.5}$$

where $a(t)$ and $b(t)$ are members of the set.

Example 6 ($C[a, b]$)**.** An inner product can be defined as an integral over the range with respect to a weight function $w(x)$:

$$\langle\, f \,|\, g \,\rangle = \int_a^b f^*(x)\, g(x)\, w(x)\, \mathrm{d}x \,. \tag{1.6}$$

Definition 2. A basis set $|\, e_n \,\rangle$, $n = 1, \ldots, N$ is **orthonormal** if

$$\langle\, e_i \,|\, e_j \,\rangle = \delta_{ij} \,.$$

We first turn to a property of the inner product: the Schwartz inequality.

Theorem 2 (The Schwartz inequality)**.** *The Schwartz, or triangle, inequality states that for any two vectors in $\mathcal{V}$,*

$$\|\psi\|\|\phi\| \geq |\langle\, \psi \,|\, \phi \,\rangle| \,.$$

Proof. We let

$$|\, \chi \,\rangle = |\, \psi \,\rangle + \lambda |\, \phi \,\rangle \,.$$

Then the length of $|\, \chi \,\rangle$ is positive definite:

$$\|\chi\|^2 = \langle\, \chi \,|\, \chi \,\rangle = \langle\, \psi \,|\, \psi \,\rangle + \lambda \langle\, \psi \,|\, \phi \,\rangle + \lambda^* \langle\, \phi \,|\, \psi \,\rangle + |\lambda|^2 \langle\, \phi \,|\, \phi \,\rangle \geq 0 \,. \tag{1.7}$$

This expression, as a function of λ and λ^*, will be a minimum when

$$\begin{aligned} \frac{\partial \langle\, \chi \,|\, \chi \,\rangle}{\partial \lambda} &= \langle\, \psi \,|\, \phi \,\rangle + \lambda^* \langle\, \phi \,|\, \phi \,\rangle = 0 \,, \\ \frac{\partial \langle\, \chi \,|\, \chi \,\rangle}{\partial \lambda^*} &= \langle\, \phi \,|\, \psi \,\rangle + \lambda \langle\, \phi \,|\, \phi \,\rangle \;\; = 0 \,. \end{aligned}$$

Thus

$$\lambda = -\langle\, \phi \,|\, \psi \,\rangle / \, \|\phi\|^2 \,, \qquad \lambda^* = -\langle\, \psi \,|\, \phi \,\rangle / \, \|\phi\|^2 \,.$$

Substituting this into (1.7) and taking the square root, we find:

$$\|\psi\|\|\phi\| \geq |\langle\, \psi \,|\, \phi \,\rangle| \,.$$

□

The Schwartz inequality allows us to generalize the idea of an "angle" between two vectors. If we let

$$\cos\gamma = |\langle\, \psi \,|\, \phi \,\rangle| / (\, \|\psi\|\|\phi\| \,) \,,$$

then the inequality states that $0 \leq \cos\gamma \leq 1$.

1.3.1 The dual space

The dual "vector" is not a vector at all, but a function which operates on vectors to produce complex numbers defined by the inner product. The dual $\langle\, \alpha \,|$ is written with a "slot" ( ) for the vectors:

$$\langle\, \alpha \,|(\quad) = g(\,|\, \alpha \,\rangle, \quad) \,, \tag{1.8}$$

for all vectors $|\, \alpha \,\rangle$ in $\mathcal{V}$. That is, the dual only makes sense if it is acting on an arbitrary vector in $\mathcal{V}$ to produce a number:

$$\langle\, \alpha \,|(\,|\, \beta \,\rangle) = g(\,|\, \alpha \,\rangle, |\, \beta \,\rangle) \equiv \langle\, \alpha \,|\, \beta \,\rangle \,, \tag{1.9}$$

in agreement with our notation for inner product. The anti-linear property of the first slot of the inner product means that the set of dual functions form an anti-linear vector space also, called $\mathcal{V}_D$. So if we regard the dual $\langle\, \alpha \,|$ as *right* acting, we can just omit the parenthesis and the slot when writing the dual function.

So if the set $|\, e_i \,\rangle$, $i = 1, \ldots, N$ are a basis in $\mathcal{V}$, then the duals of the basis vectors are defined by:

$$\langle\, e_i \,| = g(\,|\, e_i \,\rangle, \quad) \,, \tag{1.10}$$

with the property that $\langle e_i \,|\, e_j \rangle = g_{ij}$. Then if $|\, \alpha \,\rangle$ is a vector in $\mathcal{V}$ with the expansion:

$$|\, \alpha \,\rangle = \sum_i a_i \,|\, e_i \,\rangle \,. \tag{1.11}$$

Because of the anti-linear properties of the first slot in the definition of the inner product, the dual $\langle \alpha \,|$ is given uniquely by:

$$\langle \alpha \,| = \sum_i a_i^* \,\langle e_i \,| \,. \tag{1.12}$$

1.3.2 Non-orthogonal basis sets

If the basis set we have found for a linear vector space is not orthogonal, we have two choices: we can either construct an orthogonal set from the linearly independent basis set, or introduce contra- and covariant vectors. We first turn to the Gram-Schmidt orthogonalization method.

Gram-Schmidt orthogonalization

Given an arbitrary basis set, $|\, x_n \,\rangle$, $n = 1, \ldots, N$, we can construct an orthonormal basis $|\, e_n \,\rangle$, $n = 1, \ldots, N$ as follows:

1. Start with $|\, e_1 \,\rangle = |\, x_1 \,\rangle / \|x_1\|$.

2. Next construct a vector orthogonal to $|\, e_1 \,\rangle$ from $|\, x_2 \,\rangle$ and $|\, e_1 \,\rangle$, and normalize it:

$$|\, e_2 \,\rangle = \frac{|\, x_2 \,\rangle - |\, e_1 \,\rangle\langle\, e_1 \,|\, x_2 \,\rangle}{\| \,|\, x_2 \,\rangle - |\, e_1 \,\rangle\langle\, e_1 \,|\, x_2 \,\rangle \,\|} \,.$$

3. Generalize this formula to the remaining vectors:

$$|\, e_n \,\rangle = \frac{|\, x_n \,\rangle - \sum_{m=1}^{n-1} |\, e_m \,\rangle\langle\, e_m \,|\, x_n \,\rangle}{\| \,|\, x_n \,\rangle - \sum_{m=1}^{n-1} |\, e_m \,\rangle\langle\, e_m \,|\, x_n \,\rangle \,\|} \,.$$

   for $n = 2, \ldots, N$.

Contra- and co-variant vectors

Another method of dealing with non-orthogonal basis vectors is to introduce contra- and co-variant vectors. We do that in this section. We first introduce a "metric" tensor g_{ij} by the definition:

$$g_{ij} \equiv g(\,|\, e_i \,\rangle, |\, e_j \,\rangle\,) = \langle\, e_i \,|\, e_j \,\rangle \,, \tag{1.13}$$

and assume that $\det\{g\} \neq 0$, so that the inverse metric, which we write with upper indices $g_{ij}^{-1} \equiv g^{ij}$ exists:

$$\sum_j g_{ij}\, g^{jk} = \sum_j g^{ij}\, g_{jk} = \delta_{ik} \,, \tag{1.14}$$

We sometimes write: $g^i{}_k \equiv \delta_{ik}$.

Definition 3 (covariant vectors)**.** We call the basis vectors $|\, e_i \,\rangle$ with *lower* indices *co*-variant vectors,[2] and then define basis vectors $|\, e^i \,\rangle$ with upper indices by:

$$|\, e^i \,\rangle = \sum_j |\, e_j \,\rangle\, g^{ji} \,. \tag{1.15}$$

We call these vectors with upper indices *contra*-variant vectors. The duals of the *contra*-variant vectors are then given by:

$$\langle\, e^i \,| = \sum_j [\, g^{ji} \,]^* \,\langle\, e_j \,| = \sum_j g^{ij} \,\langle\, e_j \,| \,. \tag{1.16}$$

Remark 1. It is easy to show that the contra- and co-variant vectors obey the relations:

$$\langle\, e^i \,|\, e_j \,\rangle = \langle\, e_i \,|\, e^j \,\rangle = \delta_{ij} \,, \qquad \sum_i |\, e^i \,\rangle\langle\, e_i \,| = \sum_i |\, e_i \,\rangle\langle\, e^i \,| = 1 \,. \tag{1.17}$$

The set of dual vectors $|\, e^i \,\rangle$ are not orthogonal with each other and are not normalized to one even if the set $|\, e_i \,\rangle$ is normalized to one. If the basis vectors $|\, e_i \,\rangle$ *are* orthonormal, then the contra- and co-variant basis vectors are identical, which was the case that Dirac had in mind when he invented the bra and ket notation.

Remark 2. Since the sets $|\, e_i \,\rangle$ and $|\, e^i \,\rangle$ are both complete linearly independent basis sets, although *not* orthogonal, we can write a vector in one of two ways:

$$|\, v \,\rangle = \sum_i v^i \,|\, e_i \,\rangle = \sum_i v_i \,|\, e^i \,\rangle \,, \tag{1.18}$$

from which we find:

$$\langle\, e^i \,|\, v \,\rangle = v^i \,, \qquad \text{and} \qquad \langle\, e_i \,|\, v \,\rangle = v_i \,, \tag{1.19}$$

which provides an easy methods to find the contra- and co-variant expansion coefficients of vectors. This was the reason for introducing contra- and co-variant base vectors in the first place. The two components of the vector v^i and v_i are related by the inner product matrix:

$$v_i = \sum_j g_{ij}\, v^j \,, \qquad \text{and} \qquad v^i = \sum_j g^{ij}\, v_j \,. \tag{1.20}$$

That is, g_{ij} and g^{ij} "lower" and "raise" indices respectively. We now turn to a few examples.

Example 7 ($\mathbb{C}^2$)**.** Let us consider example 1 with $N = 2$, a two dimensional vector space of complex numbers. Vectors in this space are called "spinors." A vector $|\, a \,\rangle$ is written as a two-component column matrix:

$$|\, a \,\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \,. \tag{1.21}$$

The inner product of two vectors $|\, a \,\rangle$ and $|\, b \,\rangle$ is given by definition (1.4):

$$\langle\, a \,|\, b \,\rangle = a_1^* b_1 + a_2^* b_2 \,. \tag{1.22}$$

Let us now take two linearly independent non-orthogonal basis vectors given by:

$$|\, e_1 \,\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \,, \qquad |\, e_2 \,\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 1 \end{pmatrix} \,. \tag{1.23}$$

The basis bra's $\langle\, e_1 \,|$ and $\langle\, e_2 \,|$ are then given by:

$$\langle\, e_1 \,| = \begin{pmatrix} 1, & 0 \end{pmatrix} \,, \qquad \langle\, e_2 \,| = \frac{1}{\sqrt{2}} \begin{pmatrix} -i, & 1 \end{pmatrix} \,. \tag{1.24}$$

[2] Sometimes the co-variant vectors are called *dual* vectors. We do not use that terminology here because of the confusion with our definition of dual operators and the dual space of bra's.

So the g_{ij} matrix is given by:

$$g_{ij} = \begin{pmatrix} 1 & i/\sqrt{2} \\ -i/\sqrt{2} & 1 \end{pmatrix} , \tag{1.25}$$

with $\det\{g\} = 1/2$. The inverse g^{ij} is then:

$$g^{ij} = \begin{pmatrix} 2 & -i\sqrt{2} \\ i\sqrt{2} & 2 \end{pmatrix} . \tag{1.26}$$

So the contra-variant vectors are given by:

$$|\, e^1 \,\rangle = \sum_{i=1}^{2} |\, e_i \,\rangle\, g^{i1} = \begin{pmatrix} 1 \\ i \end{pmatrix} , \qquad |\, e^2 \,\rangle = \sum_{i=1}^{2} |\, e_i \,\rangle\, g^{i2} = \begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix} , \tag{1.27}$$

with the duals:

$$\langle\, e^1 \,| = \begin{pmatrix} 1, & -i \end{pmatrix} , \qquad \langle\, e^2 \,| = \begin{pmatrix} 0, & \sqrt{2} \end{pmatrix} . \tag{1.28}$$

It is easy to see that these contra- and co-variant vectors satisfy the relations:

$$\begin{aligned} \langle\, e^i \,|\, e_j \,\rangle &= \delta_{ij} , \qquad \text{and} \qquad |\, e^1 \,\rangle\langle\, e_1 \,| + |\, e^2 \,\rangle\langle\, e_2 \,| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \\ \langle\, e_i \,|\, e^j \,\rangle &= \delta_{ij} , \qquad \text{and} \qquad |\, e_1 \,\rangle\langle\, e^1 \,| + |\, e_2 \,\rangle\langle\, e^2 \,| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \end{aligned} \tag{1.29}$$

Notice that the dual vectors are *not normalized* nor are they orthogonal with each other. They *are* orthonormal with the base vectors, however, which is the important requirement.

Example 8 (P^∞)**.** In this example, we take the non-orthogonal basis set to be powers of x. So we define *co*-variant vectors $|\, e_i \,\rangle$ by:

$$|\, e_i \,\rangle = x^i , \qquad \text{for } i = 0, 1, 2, \ldots, \infty, \tag{1.30}$$

with an inner product rule given by integration over the range $[-1, 1]$:

$$\begin{aligned} g_{ij} = \langle\, e_i \,|\, e_j \,\rangle &= \int_{-1}^{+1} x^i\, x^j \,\mathrm{d}x = \begin{cases} 0 & \text{for } i+j \text{ odd}, \\ 2/(i+j+1) & \text{for } i+j \text{ even}. \end{cases} \\ &= \begin{pmatrix} 2 & 0 & 2/3 & 0 & \cdots \\ 0 & 2/3 & 0 & 2/5 & \cdots \\ 2/3 & 0 & 2/5 & 0 & \cdots \\ 0 & 2/5 & 0 & 2/7 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} . \end{aligned} \tag{1.31}$$

We would have to invert this matrix to find g^{ij}. This is not easy to do, and is left to the reader.

1.4 Operators

An operator S maps a vector $|\, \alpha \,\rangle \in \mathcal{V}$ to another vector $|\, \beta \,\rangle \in \mathcal{V}$. We write: $S(|\, \alpha \,\rangle) = |\, \beta \,\rangle$, which is defined for some domain D and range R of vectors in the space. We usually consider cases where the domain and range is the full space $\mathcal{V}$.

Observables in quantum theory are represented by Hermitian operators and symmetry transformations by unitary or anti-unitary operators. These important operators are defined in this section for finite vector spaces. We start with a number of useful definitions.

Definition 4 (linear and anti-linear operators)**.** A **linear operator** $S = L$ has the properties:

$$L(a\,|\,\alpha\,\rangle + b\,|\,\beta\,\rangle\,) = a\,L(|\,\alpha\,\rangle) + b\,L(|\,\beta\,\rangle)\,, \tag{1.32}$$

for any $|\,\alpha\,\rangle, |\,\beta\,\rangle \in \mathcal{V}$ and $a, b \in \mathbb{C}$. Similarly, an **anti-linear operator** $S = A$ has the properties:

$$A(a\,|\,\alpha\,\rangle + b\,|\,\beta\,\rangle\,) = a^*\,A(|\,\alpha\,\rangle) + b^*\,A(|\,\beta\,\rangle)\,. \tag{1.33}$$

So, for linear and anti-linear operators, we can just write: $L|\,\alpha\,\rangle = |\,L\,\alpha\,\rangle = |\,\gamma\,\rangle$, and $A|\,\alpha\,\rangle = |\,A\,\alpha\,\rangle = |\,\gamma\,\rangle$, without the parenthesis.

Definition 5 (inverse operators)**.** If the mapping $S\,|\,\alpha\,\rangle = |\,\beta\,\rangle$ is such that each $|\,\beta\,\rangle$ comes from a unique $|\,\alpha\,\rangle$, the mapping is called **injective**. In addition, if every vector $|\,\beta\,\rangle \in \mathcal{V}$ is of the form $|\,\beta\,\rangle = S\,|\,\alpha\,\rangle$, then the mapping is called **surjective**. If the mapping is both injective and surjective it is called **bijective**, and the inverse exists. (Physicists usually don't use such fancy names.) We write the **inverse** operation thus: $S^{-1}\,|\,\beta\,\rangle = |\,\alpha\,\rangle$. Clearly, $S\,S^{-1} = S^{-1}\,S = 1$, the unit operator. A bijective, linear mapping is called an **isomorphism**.

Remark 3. We note the following two theorems, which we state without proof:[3]

1. If L is a linear operator, then so is L^{-1}.
2. A linear operator has an inverse if and only if $L\,|\,\alpha\,\rangle = |\,0\,\rangle$ implies that $|\,\alpha\,\rangle = |\,0\,\rangle$.

Definition 6 (adjoint operators)**.** For **linear operators**, the **adjoint** operator $L^\dagger$ is defined by:

$$\langle\,\alpha\,|\,L^\dagger\,\beta\,\rangle = \langle\,L\,\alpha\,|\,\beta\,\rangle\,. \tag{1.34}$$

For **anti-linear** operators, the **adjoint** $A^\dagger$ is defined by:

$$\langle\,\alpha\,|\,A^\dagger\,\beta\,\rangle = \langle\,A\,\alpha\,|\,\beta\,\rangle^* = \langle\,\beta\,|\,A\,\alpha\,\rangle\,. \tag{1.35}$$

Remark 4. For an orthonormal basis, the adjoint matrix of a linear operator is

$$L^\dagger_{ij} = \langle\,e_i\,|\,L^\dagger\,e_j\,\rangle = \langle\,L\,e_i\,|\,e_j\,\rangle = \langle\,e_j\,|\,L\,e_i\,\rangle^* = L^*_{ji}\,,$$

which is the complex conjugate of the transpose matrix. For an anti-linear operator, the adjoint matrix is given by:

$$A^\dagger_{ij} = \langle\,e_i\,|\,A^\dagger\,e_j\,\rangle = \langle\,A\,e_i\,|\,e_j\,\rangle^* = \langle\,e_j\,|\,A\,e_i\,\rangle = A_{ji}\,,$$

which is the transpose matrix.

Definition 7 (unitary operators)**.** A **linear** operator U is **unitary** if

$$\langle\,U\,\alpha\,|\,U\,\beta\,\rangle = \langle\,\alpha\,|\,U^\dagger U\,\beta\,\rangle = \langle\,\alpha\,|\,\beta\,\rangle\,. \tag{1.36}$$

An **anti-linear** and **anti-unitary** operator U is defined by:

$$\langle\,U\,\alpha\,|\,U\,\beta\,\rangle = \langle\,\alpha\,|\,U^\dagger U\,\beta\,\rangle^* = \langle\,\alpha\,|\,\beta\,\rangle^* = \langle\,\beta\,|\,\alpha\,\rangle\,. \tag{1.37}$$

Thus for both unitary and anti-unitary operators $U^{-1} = U^\dagger$. This was the reason for our differing definitions of the adjoint for linear and anti-linear operators.

Definition 8 (Hermitian operators)**.** A **linear** operator H is **Hermitian** if

$$\langle\,\alpha\,|\,H\,\beta\,\rangle = \langle\,H\,\alpha\,|\,\beta\,\rangle = \langle\,\alpha\,|\,H^\dagger\,\beta\,\rangle\,. \tag{1.38}$$

That is $H^\dagger = H$.

[3] the proofs can be found in Serot[1]

Remark 5. By convention, operators are **right** acting on kets: $S\,|\,\alpha\,\rangle = |\,\beta\,\rangle$. The product of two linear operators is defined by:

$$(AB)\,|\,\alpha\,\rangle = A\,(B\,|\,\alpha\,\rangle) = A\,|\,\beta\,\rangle = |\,\gamma\,\rangle\,, \tag{1.39}$$

which is called the *composition* law, whereas

$$(BA)\,|\,\alpha\,\rangle = B\,(A\,|\,\alpha\,\rangle) = B\,|\,\delta\,\rangle = |\,\epsilon\,\rangle\,. \tag{1.40}$$

Thus in general $AB \neq BA$. Linear operators obey:

1. $A(BC) = (AB)C$; *associative law*
2. $(A+B)C = AC + BC$; *distributive law*
   $A(B+C) = AB + AC$.

The difference between the order of operations are important in quantum mechanics — we call this difference, the commutator, and write: $[A, B] = AB - BA$. In appendix B, we list some operator identities for commutators.

Remark 6. **left** acting operators act in the space of linear functionals, or dual space $\mathcal{V}_D$ of bra vectors, and are defined by the relation:

$$\begin{aligned}(\,\langle\,\alpha\,|\,S\,)\,|\,\beta\,\rangle \equiv \langle\,\alpha\,|\,(\,S\,|\,\beta\,\rangle) &= \langle\,\alpha\,|\,S\,\beta\,\rangle \\ &= \begin{cases}\langle\,S^\dagger\,\alpha\,|\,\beta\,\rangle & \text{for linear operators,}\\ \langle\,S^\dagger\,\alpha\,|\,\beta\,\rangle^* & \text{for anti-linear operators.}\end{cases}\end{aligned} \tag{1.41}$$

Thus, if for linear operators in $\mathcal{V}$ we have the relation

$$L\,|\,\alpha\,\rangle = |\,L\,\alpha\,\rangle = |\,\beta\,\rangle\,, \tag{1.42}$$

then the corresponding mapping in $\mathcal{V}_D$ is given by:

$$\langle\,\alpha\,|\,L^\dagger = \langle\,L\,\alpha\,| = \langle\,\beta\,|\,. \tag{1.43}$$

Dirac invented a notation for this. For an arbitrary operator S, he defined:

$$\langle\,\alpha\,|\,S\,|\,\beta\,\rangle \equiv \langle\,\alpha\,|\,S\,\beta\,\rangle = \begin{cases}\langle\,S^\dagger\,\alpha\,|\,\beta\,\rangle & \text{for linear operators,}\\ \langle\,S^\dagger\,\alpha\,|\,\beta\,\rangle^* & \text{for anti-linear operators.}\end{cases} \tag{1.44}$$

We often call $\langle\,\alpha\,|\,S\,|\,\beta\,\rangle$ a *matrix element*, the reasons for which will become apparent in the next section. In Dirac's notation, we can think of S as right acting on a ket vector or left acting on a bra vector.

Definition 9. A **Normal** operator is one that commutes with it's adjoint: $[A, A^\dagger] = 0$.

Remark 7. We shall learn below that normal operators are the most general kind of operators that can be diagonalized by a unitary transformation. Both unitary and Hermitian operators are examples of normal operators.

1.4.1 Eigenvalues and eigenvectors:

For *any* operator A, if we can find a complex number a and a ket $|\,a\,\rangle$ such that

$$A|\,a\,\rangle = a|\,a\,\rangle\,, \tag{1.45}$$

then a is called the eigenvalue and $|\,a\,\rangle$ the eigenvector. We assume in this section that the domain and range of the operator A is the full set $\mathcal{V}$. Note that we have simplified our notation here by labeling the eigenvector by the eigenvalue a, rather than using Greek characters for labeling vectors. Vectors are distinguished by Dirac's ket notation.

Theorem 3 (Hermitian operators)**.** *The eigenvalues of Hermitian operators are real and the eigenvectors can be made orthonormal.*

Proof. The eigenvalue equation for Hermitian operators is written:

$$H\,|\,h\,\rangle = h\,|\,h\,\rangle \tag{1.46}$$

Since H is Hermitian, we have the following relations:

$$\langle\, h'\,|\,H\,h\,\rangle = h\,\langle\, h'\,|\,h\,\rangle\,,$$
$$\langle\, h'\,|\,H^\dagger\,h\,\rangle = \langle\, H\,h'\,|\,h\,\rangle = h'^*\,\langle\, h'\,|\,h\,\rangle\,.$$

But for hermitian operators $H^\dagger = H$, so subtracting these two equations, we find:

$$(h - h'^*)\,\langle\, h'\,|\,h\,\rangle = 0\,. \tag{1.47}$$

So setting $h' = h$, we have

$$(h - h^*)\,\|h\| = 0\,.$$

Since $\|h\| \neq 0$, $h = h^*$. Thus h is real. Since all the eigenvalues are real, we have from (1.47),

$$(h - h')\,\langle\, h'\,|\,h\,\rangle = 0\,. \tag{1.48}$$

Thus, if $h \neq h'$, then $\langle\, h'\,|\,h\,\rangle = 0$, and is orthogonal. The proof of orthogonality of the eigenvectors fails if there is more that one eigenvector with the same eigenvalue (we call this degenerate eigenvalues). However, by the Gram-Schmidt construction, that we can always find orthogonal eigenvectors among the eigenvectors with degenerate eigenvalues. Since $\|h\| \neq 0$, it is always possible to normalize them. Thus we can assume that hermitian operators have real and ortho*normal* eigenvectors, $\langle\, h'\,|\,h\,\rangle = \delta_{h'h}$. □

Theorem 4 (Unitary operators)**.** *The eigenvalues of unitary operators have unit magnitude and the eigenvectors can be made orthogonal.*

Proof. The eigenvalue equation for Unitary operators is written:

$$U\,|\,u\,\rangle = u\,|\,u\,\rangle \tag{1.49}$$

Unitary operators obey $U^\dagger U = 1$, so we have:

$$\langle\, u'\,|\,U^\dagger U\,u\,\rangle = \langle\, U\,u'\,|\,U\,u\,\rangle = u'^* u\,\langle\, u'\,|\,u\,\rangle = \langle\, u'\,|\,u\,\rangle\,.$$

So we find:

$$(1 - u'^* u)\langle\, u'\,|\,u\,\rangle = 0\,.$$

Therefore, if $u' = u$, $(1 - |u|^2)\|u\| = 0$, and we must have $|u| = 1$. This means we can write $u = e^{i\theta}$, where θ is real. In addition if $u \neq u'$, then the eigenvectors are orthogonal, $\langle\, u'\,|\,u\,\rangle = 0$. Degenerate eigenvalues can again be orthonormalized by the Gram-Schmidt method. □

Remark 8 (finding eigenvalues and eigenvectors). For finite systems, we can find eigenvalues and eigenvectors of an operator A by solving a set of linear equations. Let $|\,e_i\,\rangle$, $i = 1, \ldots, N$ be an orthonormal basis in $\mathcal{V}$. Then we can write:

$$|\,a\,\rangle = \sum_{i=1}^{N} |\,e_i\,\rangle\,c_i(a)\,, \qquad c_i(\alpha) = \langle\, e_i\,|\,a\,\rangle\,. \tag{1.50}$$

Then Eq. (1.45), becomes:

$$\sum_{j=1}^{N} A_{ij}\,c_j(a) = a\,c_i(a)\,, \qquad \text{for } i = 1, \ldots, N, \tag{1.51}$$

where $A_{ij} = \langle e_i \,|\, A \,|\, e_j \rangle$. We write $\underline{A}$ as the *matrix* with elements A_{ij}. By Cramer's rule[2][p. 92], Eq. (1.51) has nontrivial solutions if

$$f(a) = \det[\,\underline{A} - a\underline{I}\,] = 0\,. \tag{1.52}$$

$f(a)$ is a polynomial in a of degree N, and is called the **characteristic** polynomial. In general, it has N complex roots, some of which may be the same. We call the number of multiple roots the **degeneracy** of the root. If we call all these roots a_n then we can write formally:

$$f(a) = \sum_{n=1}^{N} c_n\, a^n = \prod_{n=1}^{N} (a - a_n) = 0\,, \tag{1.53}$$

For Hermitian operators, these roots are all real. For unitary operators, they are complex with unit magnitude. The coefficients $c_i(a)$ can be found for each eigenvalue from the linear set of equations (1.51). If there are no multiple roots, the N eigenvectors so found are orthogonal and span $\mathcal{V}$. For the case of multiple roots, it is possible to still find N linearly independent eigenvectors and orthorgonalize them by the Schmidt procedure. Thus we can assume that we can construct, by these methods, a complete set of orthonormal vectors that span the vector space.

1.4.2 Non-orthogonal basis vectors

Let us examine in this section how to write matrix elements of operators using non-orthogonal basis sets. First let A be a linear operator satisfying:

$$A \,|\, v \,\rangle = |\, u \,\rangle \tag{1.54}$$

Expanding the vectors in terms of the co-variant basis set $|\, e_i \,\rangle$, we have:

$$A \,|\, v \,\rangle = \sum_j v^j\, A \,|\, e_j \,\rangle = |\, u \,\rangle = \sum_i u^i \,|\, e_i \,\rangle\,, \tag{1.55}$$

Right operating on this by the bra $\langle\, e^i \,|$ gives:

$$\sum_j A^i{}_j\, v^j = u^i\,, \qquad \text{where} \qquad A^i{}_j = \langle\, e^i \,|\, A\, e_j \,\rangle \equiv \langle\, e^i \,|\, A \,|\, e_j \,\rangle\,. \tag{1.56}$$

This can be interpreted as matrix multiplication of a square matrix $A^i{}_j$ with the column matrix v^j to give the column matrix u^i. Similarly, expanding the vectors in terms of the contra-variant basis vectors $|\, e^i \,\rangle$ gives a corresponding expression:

$$A \,|\, v \,\rangle = \sum_j v_j\, A \,|\, e^j \,\rangle = |\, u \,\rangle = \sum_i u_i \,|\, e^i \,\rangle\,, \tag{1.57}$$

Right operating again on this expression by $\langle\, e_i \,|$ gives:

$$\sum_j A_i{}^j\, v_j = u_i\,, \qquad \text{where} \qquad A_i{}^j = \langle\, e_i \,|\, A\, e^j \,\rangle \equiv \langle\, e_i \,|\, A \,|\, e^j \,\rangle\,. \tag{1.58}$$

This can also be interpreted as matrix multiplication of a square matrix $A_i{}^j$ with the column matrix v_j to give the column matrix u_i. One can easily check that

$$A_i{}^j = \sum_{i'j'} g_{ii'}\, A^{i'}{}_{j'}\, g^{j'j}\,. \tag{1.59}$$

$A_i{}^j$ and $A^i{}_j$ are *not* the same matrix. We can also define matrices A_{ij} and A^{ij} by:

$$A_{ij} = \langle\, e_i \,|\, A \,|\, e_j \,\rangle\,, \qquad \text{and} \qquad A^{ij} = \langle\, e^i \,|\, A \,|\, e^j \,\rangle\,. \tag{1.60}$$

These matrices are related to the others by raising or lowering indices using $g_{ii'}$ and $g^{jj'}$. For example, we have:

$$A_{ij} = \sum_{i'} g_{ii'} A^{i'}{}_{j} = \sum_{j'} A_i{}^{j'} g_{j'j} = \sum_{i'j'} g_{ii'} A^{i'j'} g_{j'j} \,. \tag{1.61}$$

Example 9 (Adjoint)**.** From definition (6) of the adjoint of linear operators, we find that matrix elements in non-orthogonal basis sets are given by:

$$\begin{aligned} [L^\dagger]_{ij} &= [L_{ji}]^* \,, & [L^\dagger]^i{}_j &= [L_i{}^j]^* \\ [L^\dagger]_i{}^j &= [L^i{}_j]^* & [L^\dagger]^{ij} &= [L^{ji}]^* \,. \end{aligned} \tag{1.62}$$

Only the first and last of these matrix forms give a definition of adjoint matrix that relate the complex conjugate of a matrix to the transpose of the same matrix. The second and third form relate complex conjugates of matrices with upper and lower indices to the transpose of matrices with lower and upper indices, and do not even refer to the same matrix.

Example 10 (Hermitian operators)**.** For Hermitian operators such that $\hat{H}^\dagger = \hat{H}$, we find:

$$\begin{aligned} H_{ij} &= [H_{ji}]^* \,, & H^i{}_j &= [H_i{}^j]^* \\ H_i{}^j &= [H^i{}_j]^* & H^{ij} &= [H^{ji}]^* \,. \end{aligned} \tag{1.63}$$

Example 11 (P^∞)**.** Returning to our example of continuous functions defined on the interval $[-1, 1]$ and using the non-orthogonal basis P^∞, as defined in Example 8 above, we define an operator X which is multiplication of functions in the vector space by x. We take the *co*-variant basis vectors to be given by Eq. (1.30). On this basis set, the X operation is easily described as:

$$X \,|\, e_i \,\rangle = |\, e_{i+1} \,\rangle \,, \qquad \text{for } i = 0, 1, \ldots, \infty. \tag{1.64}$$

So we easily construct the mixed tensor $X^i{}_j$:

$$X^i{}_j = \langle\, e^i \,|\, X \,|\, e_j \,\rangle = \langle\, e^i \,|\, e_{j+1} \,\rangle = \delta_{i,j+1} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} . \tag{1.65}$$

However, there are three other tensors we can write down. For example, the matrix X_{ij} is given by:

$$X_{ij} = \langle\, e_i \,|\, X \,|\, e_j \,\rangle = \langle\, e_i \,|\, e_{j+1} \,\rangle = g_{i,j+1} = \begin{pmatrix} 0 & 2/3 & 0 & 2/5 & \cdots \\ 2/3 & 0 & 2/5 & 0 & \cdots \\ 0 & 2/5 & 0 & 2/7 & \cdots \\ 2/5 & 0 & 2/7 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} . \tag{1.66}$$

This matrix obeys $X^*_{ji} = X_{ij}$, and is clearly Hermitian. Since we (now) know that X *is* a hermitian operator, we can deduce that:

$$X_i{}^j = [\, X^i{}_j \,]^* = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} , \tag{1.67}$$

without knowing what the *contra*-variant vectors are.

1.4.3 Projection operators:

Divide the vector space into two parts: $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$. (This is called the **direct sum**.) Then any vector $|\,a\,\rangle$ can be written as a sum of two vectors $|\,a\,\rangle = |\,a_1\,\rangle + |\,a_2\,\rangle$, where $|\,a_1\,\rangle \in \mathcal{V}_1$ and $|\,a_2\,\rangle \in \mathcal{V}_2$. Then we define projection operators P_1 and P_2 by their action on an arbitrary vector $|\,a\,\rangle \in \mathcal{V}$:

$$\begin{aligned} P_1\,|\,a\,\rangle &= |\,a_1\,\rangle\,, \qquad |\,a_1\,\rangle \in \mathcal{V}_1\,, \\ P_2\,|\,a\,\rangle &= |\,a_2\,\rangle\,, \qquad |\,a_2\,\rangle \in \mathcal{V}_2\,, \end{aligned} \tag{1.68}$$

Obviously, $P_1 + P_2 = 1$, and $P_1^2 = P_1$, $P_1^\dagger = P_1$, with similar relations for P_2. We can continue to divide the vector space into a maximum of N divisions. In fact, let $|\,e_i\,\rangle$ for $i = 1, \ldots, N$ be an orthonormal basis for $\mathcal{V}$. Then

$$P_i = |\,e_i\,\rangle\langle\,e_i\,|\,, \qquad P_iP_j = \delta_{ij}P_i\,, \qquad \sum_{i=1}^{N} P_i = 1\,, \tag{1.69}$$

with $P_i \neq 0$, divides the vector space up into a direct sum of *one*-dimensional parts:

$$\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_N\,.$$

The action of P_i on an arbitrary vector $|\,b\,\rangle$ gives

$$P_i\,|\,b\,\rangle = |\,e_i\,\rangle\,\langle\,e_i\,|\,b\,\rangle\,.$$

1.4.4 Spectral representations:

Let us now write the eigenvalue equation (1.45) for a normal operator in the following way:

$$A|\,a_{ij}\,\rangle = a_i\,|\,a_{ij}\,\rangle\,, \tag{1.70}$$

$$\text{where} \quad i = 1, \ldots, n, \qquad \text{with } n \leq N,$$

$$\text{and} \quad j = 1, \ldots, m_i. \tag{1.71}$$

Here a_i are the m_i-fold degenerate eigenvalues of A. From our discussion in Section 1.4.1, we conclude that these eigenvectors are all orthonormal:

$$\langle\,a_{ij}\,|\,a_{i'j'}\,\rangle = \delta_{i,i'}\,\delta_{j,j'}\,, \tag{1.72}$$

and span the space. Thus for *any* vector $|\,b\,\rangle$, we can write:

$$|\,b\,\rangle = \sum_{i=1}^{n}\sum_{j=1}^{m_i} |\,a_{ij}\,\rangle\,c_{ij}(b)\,, \qquad \text{where} \qquad c_{ij}(b) = \langle\,a_{ij}\,|\,b\,\rangle\,. \tag{1.73}$$

Inserting $c_{ij}(b)$ back into the first of Eq. (1.73) gives:

$$|\,b\,\rangle = \sum_{i=1}^{n}\left\{\sum_{j=1}^{m_i} |\,a_{ij}\,\rangle\langle\,a_{ij}\,|\right\}|\,b\,\rangle = \sum_{i=1}^{n} P_i\,|\,b\,\rangle\,. \tag{1.74}$$

Here we define a projection operator by:

$$P_i = \sum_{j=1}^{m_i} |\,a_{ij}\,\rangle\langle\,a_{ij}\,|\,, \tag{1.75}$$

with the properties:

$$P_iP_j = \delta_{ij}P_i\,, \qquad \sum_{i=1}^{n} P_i = 1\,. \tag{1.76}$$

Eq. (1.76) is called the **completeness** statement. From Eq. (1.70), we find the **spectral representation** of the operator A:

$$A = \sum_{i=1}^{n} a_i \, P_i \, , \tag{1.77}$$

Matrix elements of a projection operator in an arbitrary basis $|\, e_i \,\rangle$ are defined by:

$$(\, P_k \,)_{ij} = \langle\, e_i \,|\, P_k \,|\, e_j \,\rangle = \sum_{l=1}^{m_k} \langle\, e_i \,|\, a_{kl} \,\rangle \langle\, a_{kl} \,|\, e_j \,\rangle \, .$$

The trace of the projection operator matrix is then easily found to be:

$$\mathrm{Tr}[\, P_k \,] = \sum_{i=1}^{N} \sum_{l=1}^{m_k} \langle\, e_i \,|\, a_{kl} \,\rangle \langle\, a_{kl} \,|\, e_i \,\rangle = \sum_{l=1}^{m_k} \left\{ \sum_{i=1}^{N} \langle\, a_{kl} \,|\, e_i \,\rangle \langle\, e_i \,|\, a_{kl} \,\rangle \right\} = m_k \, . \tag{1.78}$$

Now since

$$\mathrm{Tr}[\, P_i P_j \,] = \delta_{ij} \, \mathrm{Tr}[\, P_i \,] = m_i \, \delta_{ij} \, , \tag{1.79}$$

We can invert Eq. (1.77) to find the coefficient a_i in the spectral expansion:

$$a_i = \mathrm{Tr}[\, A \, P_i \,] / m_i \, . \tag{1.80}$$

The following two lemmas are given without proof.

Lemma 1. *For any power of a normal operator A, we have:*

$$A^k = \sum_{i=1}^{n} a_i^k P_i \, . \tag{1.81}$$

So for any function $f(A)$ of the operator A which we can write as a power series, we find:

$$f(A) = \sum_k f_k A^k = \sum_{i=1}^{n} \left\{ \sum_k f_k a_i^k \right\} P_i = \sum_{i=1}^{n} f(a_i) \, P_i \, .$$

Lemma 2. *For a normal operator A, we can write:*

$$f(A) = \sum_i c_k \, P_i \, , \qquad \textit{with} \qquad c_k = \mathrm{Tr}[\, f(A) \, P_i \,] / m_i \, . \tag{1.82}$$

Note that the sum in Eq. (1.82) contains only n terms. This is the minimum *number of terms in the expansion.*

We end our discussion in this section with some examples.

Example 12 (Hermitian operators)**.** A Hermitian H operator has the spectral representation:

$$H = \sum_i h_i \, P_i \, , \qquad P_i = |\, h_i \,\rangle \langle\, h_i \,| \, . \tag{1.83}$$

Example 13 (Unitary operators)**.** A unitary operator U has the spectral representation:

$$U = \sum_i u_i P_i \, , \qquad P_i = |\, u_i \,\rangle \langle\, u_i \,| \, .$$

with $u_i = e^{i\theta_i}$.

Example 14 (Inverse)**.** If A has an orthogonal and complete set of eigenvectors $|\, a_i \,\rangle$, $i = 1, \ldots, N$, then the inverse of $A - \lambda \, I$ has the spectral representation:

$$(\, A - \lambda \, I \,)^{-1} = \sum_i \frac{P_i}{a_i - \lambda} \, ,$$

for $\lambda \neq a_i$ for any i, and where $P_i = |\, a_i \,\rangle \langle\, a_i \,|$.

1.4.5 Basis transformations

Let $|\, e_i \,\rangle$ and $|\, f_i \,\rangle$, $i = 1, \ldots, N$ be two orthonormal and complete bases for $\mathcal{V}$. Then the two basis sets are related by:

$$|\, f_i \,\rangle = \sum_{j=1}^{N} |\, e_j \,\rangle\langle\, e_j \,|\, f_i \,\rangle = U|\, e_i \,\rangle\,. \tag{1.84}$$

Here U is an operator which maps $|\, e_i \,\rangle$ into $|\, f_i \,\rangle$ for all $i = 1, \ldots, N$. Multiplying (1.84) on the *right* by $\langle\, e_i \,|$ and summing over i gives:

$$U = \sum_{i=1}^{N} |\, f_i \,\rangle\langle\, e_i \,|\,. \tag{1.85}$$

The matrix elements of U are the same in both bases:

$$U_{ij} = \langle\, e_i \,|U|\, e_j \,\rangle = \langle\, f_i \,|U|\, f_j \,\rangle = \langle\, e_i \,|\, f_j \,\rangle\,. \tag{1.86}$$

The adjoint of U is:

$$U^\dagger = \sum_{i=1}^{N} |\, e_i \,\rangle\langle\, f_i \,|\,.$$

One can easily check that $U^\dagger U = UU^\dagger = 1$. Thus basis transformations are unitary transformations. Unitary transformations preserve lengths and angles. That is, if $|\, a' \,\rangle = U|\, a \,\rangle$ and $|\, b' \,\rangle = U|\, b \rangle$, then for unitary U,

$$\langle\, a' \,|\, b' \,\rangle = \langle\, a \,|U^\dagger U|\, b \,\rangle = \langle\, a \,|\, b \,\rangle\,.$$

Unitary basis transformations are the quantum theory analog of the orthogonal transformation of the basis vectors in a classical three-dimensional space coordinate systems which are related by a rotation of the coordinate system.

Example 15 (Hermitian operators)**.** We show here that Hermitian operators are diagonalized by unitary transformations. Let the eigenvalue equation for the the Hermitian operator H be given by:

$$H\,|\, h_i \,\rangle = h_i\,|\, h_i \,\rangle\,, \qquad \text{for } i = 1, \ldots, N.$$

In the spectral representation, H is given by:

$$H = \sum_{i=1}^{N} h_i\,|\, h_i \,\rangle\langle\, h_i \,|\,, \tag{1.87}$$

Now let

$$U = \sum_{i=1}^{N} |\, h_i \,\rangle\langle\, e_i \,|\,, \qquad U^\dagger = \sum_{i=1}^{N} |\, e_i \,\rangle\langle\, h_i \,|\,, \tag{1.88}$$

so that $U_{ij} = \langle\, e_i \,|\, h_j \,\rangle$. Then we define:

$$H_d = U^\dagger\, H\, U = \sum_{i,j,k=1}^{N} h_k|\, e_i \,\rangle\langle\, h_i \,|\, h_k \,\rangle\langle\, h_k \,|\, h_j \,\rangle\langle\, e_j \,| = \sum_{j=1}^{N} h_j|\, e_j \,\rangle\langle\, e_j \,|\,.$$

That is, H_d is diagonal in the original basis with the eigenvalues of H on the diagonal. H_d is not the *spectral* representation of H. Eq. (1.88) shows that the matrix elements of U in the $|\, e_j \,\rangle$ basis is made up of columns of the eigenvectors of H.

1.4.6 Commuting operators

We start out this section with the important theorem:

Theorem 5 (Commuting operators)**.** *Two Hermitian operators A and B have common eigenvectors if and only if they commute.*

Proof. First, assume that A and B have common eigenvectors, which we call $|\, c_i \,\rangle$:

$$A\,|\, c_i \,\rangle = a_i\,|\, c_i \,\rangle\,,$$
$$B\,|\, c_i \,\rangle = b_i\,|\, c_i \,\rangle\,,$$

with $i = 1, 2, \ldots, N$. Then

$$[\, A, B \,]\,|\, c_i \,\rangle = (\, a_i b_i - b_i a_i \,)\,|\, c_i \,\rangle = 0\,,$$

since a_i and b_i are numbers and commute.

Next, assume that A and B commute. Start with the basis in which B is diagonal, and let $B\,|\, b_i \,\rangle = b_i\,|\, b_i \,\rangle$, with b_i real. Then taking matrix elements of the commutation relation, we find:

$$\langle\, b_i \,|\, [\, A, B \,]\,|\, b_j \,\rangle = (b_j - b_i)\,\langle\, b_i \,|A|\, b_j \,\rangle = 0\,.$$

So if $b_i \neq b_j$, then A is diagonal in the representation in which B is diagonal. If $b_i = b_j$, then we can diagonalize A in the subspace of the degenerate eigenvalues of B without changing the eigenvalue of B. So, in this way, we can obtain common eigenvectors of *both* A and B. □

Remark 9. We denote the common eigenvectors by $|\, a, b \,\rangle$ and write:

$$A\,|\, a, b \,\rangle = a\,|\, a, b \,\rangle\,,$$
$$B\,|\, a, b \,\rangle = b\,|\, a, b \,\rangle\,.$$

Lemma 3. *For two Hermitian operators A and B with the spectral representations,*

$$A = \sum_{i=1}^{n} a_i\, P_i^{(A)}\,, \qquad \textit{and} \qquad B = \sum_{i=1}^{n} b_i\, P_i^{(B)}\,,$$

then $[\, P_i^{(A)}, P_k^{(B)} \,] = 0$ if and only if $[\, A, B \,] = 0$.

Note that A and B *must have the same degree of degeneracy.*

Proof. First of all, it is obvious that if $[\, P_i^{(A)}, P_k^{(B)} \,] = 0$, then $[\, A, B \,] = 0$. For the reverse, we find: □

Example 16 (Serot)**.** We will illustrate how to find common eigenvectors of commuting Hermitian operators by an example.[4] Consider the two matrices

$$A = \begin{pmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}\,.$$

Then

$$AB = BA = \begin{pmatrix} 9 & -9 & 0 \\ -9 & 9 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,, \qquad \text{so} \qquad [\, A, B \,] = 0\,, \tag{1.89}$$

[4] Taken from Serot [1, pp. 32–36]

and indeed these two matrices have common eigenvectors. So let us first diagonalize B. The secular equation is:

$$\begin{vmatrix} 2-\lambda & -1 & -1 \\ -1 & 2-\lambda & -1 \\ -1 & -1 & 2-\lambda \end{vmatrix} = -\lambda^3 + 6\lambda^2 - p\lambda = -\lambda(3-\lambda)(3-\lambda) = 0\,, \tag{1.90}$$

so the eigenvalues are $\lambda = 0, 3$, and 3. For $\lambda = 0$, the eigenvector equations are:

$$\begin{aligned} 2a - b - c &= 0\,, \\ -a + 2b - c &= 0\,, \\ -a - b + 2c &= 0\,, \end{aligned} \tag{1.91}$$

the solution of which is $a = b = c$, so

$$|\, b_1 \,\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\,. \tag{1.92}$$

For $\lambda = 3$, there is only one independent eigenvector equation:

$$a + b + c = 0\,, \tag{1.93}$$

which means that $c = -a - b$. Then a general solution is given by:

$$|\, b' \,\rangle = \begin{pmatrix} a \\ b \\ -a-b \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = a\,|\, b'_2 \,\rangle + b\,|\, b'_3 \,\rangle\,. \tag{1.94}$$

Now $|\, b'_2 \,\rangle$ and $|\, b'_3 \,\rangle$ are linearly independent but not orthogonal. So use Gram-Schmidt. First take

$$|\, b_2 \,\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\,, \tag{1.95}$$

which is normalized and orthogonal to $|\, b_1 \,\rangle$. Then the Gram-Schmidt procedure is to construct a vector $|\, b''_3 \,\rangle$ by writing:

$$\begin{aligned} |\, b''_3 \,\rangle &= |\, b'_3 \,\rangle - |\, b_2 \,\rangle\langle\, b_2 \,|\, b'_3 \,\rangle \\ &= \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 1 \\ -1/2 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}\,, \end{aligned} \tag{1.96}$$

which is now orthogonal to $|\, b_2 \,\rangle$ and $|\, b_1 \,\rangle$. Normalizing this vector, we find:

$$|\, b_3 \,\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}\,. \tag{1.97}$$

The phase of these vectors is arbitrary, and will not matter in the end. Putting these eigenvectors into columns in a matrix, we find that the unitary matrix:

$$U_B = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & \sqrt{3} & 1 \\ \sqrt{2} & 0 & -2 \\ \sqrt{2} & -\sqrt{3} & 1 \end{pmatrix}\,,$$

diagonalizes the matrix B. That is:

$$B_d = U_B^\dagger \, B \, U_B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} .$$

The next step is to compute the matrix A in the basis of the eigenvectors of B. That is:

$$A' = U_B^\dagger \, A \, U_B = \frac{1}{2} \begin{pmatrix} 12 & 0 & 0 \\ 0 & 3 & 3\sqrt{3} \\ 0 & 3\sqrt{3} & 9 \end{pmatrix} .$$

Note that this matrix is now block diagonal. We see by inspection that one eigenvalue of this matrix is 6 with eigenvector:

$$|\, a_1' \,\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} . \tag{1.98}$$

The other eigenvectors and eigenvalues can be found by diagonalizing the 2×2 block. We find:

$$\begin{vmatrix} 3/2 - \lambda & 3\sqrt{3}/2 \\ 3\sqrt{3}/2 & 9/2 - \lambda \end{vmatrix} = \lambda(\lambda - 6) = 0 \, . \tag{1.99}$$

The $\lambda = 0$ and 6 eigenvectors of A' are then given by:

$$|\, a_2' \,\rangle = \frac{1}{2} \begin{pmatrix} 0 \\ \sqrt{3} \\ -1 \end{pmatrix} , \qquad \text{and} \qquad |\, a_3' \,\rangle = \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ \sqrt{3} \end{pmatrix} . \tag{1.100}$$

Again putting the eigenvectors associated with these eigenvalues into a unitary matrix, we have:

$$U_{A'} = \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & \sqrt{3} & 1 \\ 0 & -1 & \sqrt{3} \end{pmatrix} .$$

Then

$$A_d = U_{A'}^\dagger \, A' \, U_{A'} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 6 \end{pmatrix} , \tag{1.101}$$

which brings A' into diagonal form. So we define

$$U = U_B \, U_{A'} = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & 1 & \sqrt{3} \\ \sqrt{2} & 1 & -\sqrt{3} \\ \sqrt{2} & -2 & 0 \end{pmatrix} , \tag{1.102}$$

the columns of which give the three eigenvectors. Now this matrix with bring *both* A and B into diagonal form. We find:

$$A_d = U^\dagger \, A \, U = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 6 \end{pmatrix} \qquad \text{and} \qquad B_d = U^\dagger \, B \, U = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} . \tag{1.103}$$

Note that *both* A and B have degenerate eigenvalues. The common eigenvectors, which we label $|\, ab \,\rangle$ where a and b are the eigenvalues of A and B, are given by:

$$|\, 6, 0 \,\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} , \qquad |\, 0, 3 \,\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} , \qquad |\, 6, 3 \,\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} ,$$

It is easy to check that these eigenvector are common eigenvector of both A and B. Note that there is no *common* eigenvector with eigenvalues $(0, 0)$, even though these are possible eigenvalues of A and B.

The spectral representation of A_d and B_d is given by:

$$A_d = 0\,P^{(A)}_{d\,0} + 6\,P^{(A)}_{d\,6}\,, \qquad B_d = 0\,P^{(B)}_{d\,0} + 3\,P^{(B)}_{d\,3}\,. \tag{1.104}$$

where the projection operators in the common diagonal basis are given by:

$$P^{(A)}_{d\,0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = |\,2\,\rangle\langle\,2\,|\,, \qquad P^{(A)}_{d\,6} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = |\,1\,\rangle\langle\,1\,| + |\,3\,\rangle\langle\,3\,|\,,$$

$$P^{(B)}_{d\,0} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = |\,1\,\rangle\langle\,1\,|\,, \qquad P^{(B)}_{d\,3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = |\,2\,\rangle\langle\,2\,| + |\,3\,\rangle\langle\,3\,|\,.$$

Products of the projection operators for A_d and B_d are given by:

$$P^{(A)}_{d\,6}\,P^{(B)}_{d\,0} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = |\,1\,\rangle\langle\,1\,| = P^{(B)}_{d\,0}\,, \qquad P^{(A)}_{d\,0}\,P^{(B)}_{d\,3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = |\,2\,\rangle\langle\,2\,| = P^{(A)}_{d\,0}\,,$$

$$P^{(A)}_{d\,6}\,P^{(B)}_{d\,3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = |\,3\,\rangle\langle\,3\,|\,, \qquad P^{(A)}_{d\,0}\,P^{(B)}_{d\,0} = 0\,.$$

Theorem 6 (Normal operators)**.** *A linear operator can be brought to diagonal form by a unitary transformation if and only if it is normal.*

Proof. We start by noting that any linear operator A can be written as: $A = B + iC$, with B and C Hermitian. Now suppose that A is normal. Then $[A, A^\dagger] = 2i[C, B] = 0$. So by Theorem 5, B and C can be simultaneously diagonalized and brought to diagonal form. Then A can be diagonalized also by the common eigenvectors of B and C.

Now suppose A can be diagonalized by a unitary transformation U. Then $U^\dagger\,A\,U = A_d$. So we find:

$$[A, A^\dagger] = [\,U\,A_d\,U^\dagger, U\,A_d^\dagger\,U^\dagger\,] = U\,[\,A_d, A_d^*\,]\,U^\dagger = 0\,,$$

since a diagonal operator always commutes with it's complex conjugate. □

We end this section with some results concerning determinates and traces. The results are true for matrices. For operators in general, the results are "symbolic" and lack mathematical rigor, but can be understood in terms of eigenvector expansions. We offer the following relations for normal operators, without proof.

Lemma 4.

$$\det[\,A\,B\,] = \det[\,B\,A\,] = \det[\,A\,]\,\det[\,B\,]\,,$$
$$\mathrm{Tr}[\,A + B\,] = \mathrm{Tr}[\,A\,] + \mathrm{Tr}[\,B\,]\,,$$
$$\mathrm{Tr}[\,A\,B\,] = \mathrm{Tr}[\,B\,A\,]\,.$$

Lemma 5. *For unitary operators,* $\det[\,U\,] = \pm 1$.

Lemma 6. *If A is a normal operator,*

$$\det[\,A\,] = \prod_i a_i\,, \qquad \mathrm{Tr}[\,A\,] = \sum_i a_i\,,$$
$$\det[\,A\,] = \exp[\,\mathrm{Tr}[\ln\{\,A\,\}]\,]\,.$$

1.4.7 Maximal sets of commuting operators

We have seen that if a given Hermitian operator A_1 in our vector space $\mathcal{V}$ has no degeneracies, then its eigenvectors span the space. However if there are degenerate eigenvalues of A_1, it is always possible to find a second operator A_2 which commutes with A_1 and which has different eigenvalues of A_2 for the degenerate states of A_1. We showed this in our example above. There could be additional degeneracies in these common eigenvalues of both A_1 and A_2, in which case it must be possible to find a third operator A_3 which has different eigenvalues for these common degenerate eigenvectors of both A_1 and A_2. Continuing in this way, we see that in general we might need a set of M commuting Hermitian operators: $A_1, A_2, \ldots, A_M$, all of which commute, and which can be used to specify uniquely the states which span $\mathcal{V}$.

In this way, we can, in principle, obtain a *maximal* set of commuting observables which span the vector space, $|\, a_1, a_2, \ldots, a_M \,\rangle$, such that

$$\langle\, a_1', a_2', \ldots, a_M' \,|\, a_1, a_2, \ldots, a_M \,\rangle = \delta_{a_1', a_1}\delta_{a_2', a_2}\cdots\delta_{a_M', a_M}\,,$$
$$\sum_{a_1, a_2 \cdots a_M} |\, a_1, a_2, \ldots, a_M \,\rangle\langle\, a_1, a_2, \ldots, a_M \,| = 1\,.$$

For any given system, it is not obvious how many observables constitute a maximal set, or how exactly to find them. We will see that degenerate eigenvalues for an operator are a result of symmetries inherent in the operators, and these symmetries may not be obvious. We study symmetries in the next chapter.

1.5 Infinite dimensional spaces

In the study of the quantum mechanics of a single particle, it is useful to have a concept of the measurement of the position or momentum of a particle. For one-dimensional quantum mechanics, this means that we should define Hermitian operators X and P in our vector space which have continuous real eigenvalues. We write these eigenvalue equations as:

$$X\,|\, x \,\rangle = x\,|\, x \,\rangle\,, \qquad -\infty \le x \le +\infty\,, \tag{1.105}$$
$$P\,|\, p \,\rangle = p\,|\, p \,\rangle\,, \qquad -\infty \le p \le +\infty\,. \tag{1.106}$$

There are an infinite number of these basis vectors.[5] The inner product of the coordinate and momentum vectors are defined using Dirac delta-functions. We take them to be:

$$\langle\, x \,|\, x' \,\rangle = \delta(x - x')\,, \tag{1.107}$$
$$\langle\, p \,|\, p' \,\rangle = (2\pi\hbar)\,\delta(p - p')\,. \tag{1.108}$$

Note that these vectors are not normalized, but in fact: $\|x\|^2 = \langle\, x \,|\, x \,\rangle = \infty$. This violates one of our basic assumptions about the property of the inner product, and requires care in dealing with such concepts as traces and determinants of operators. However, it is common practice to relax the normalization requirement of the inner product to include this kind of Dirac delta-function normalization. When we do so, the vector space is called "rigged."

When we study the canonical quantization methods in a later chapter, we will find that X and P, if they are to describe the position and momentum of a single particle, do not commute, and so constitute *alternative* descriptions of the particle. In fact, we will find that:

$$[\, X, P \,] = i\,\hbar\,. \tag{1.109}$$

Thus the kets $|\, x \,\rangle$ and $|\, p \,\rangle$ are two different basis sets for a vector space describing a single particle. So let the vector $|\, \psi \,\rangle$ describe the state of the particle. Then we can expand $|\, \psi \,\rangle$ in terms of either of these two basis vectors:

$$|\, \psi \,\rangle = \int \mathrm{d}x\,|\, x \,\rangle\,\psi(x) = \int \frac{\mathrm{d}p}{(2\pi\hbar)}\,|\, p \,\rangle\,\tilde{\psi}(p)\,, \tag{1.110}$$

[5] x and p have units also!

where

$$\psi(x) = \langle\, x \,|\, \psi \,\rangle\,, \qquad \tilde{\psi}(p) = \langle\, p \,|\, \psi \,\rangle\,. \tag{1.111}$$

Since in quantum mechanics, we interpret the probability of finding the particle somewhere as the length of the vector $|\,\psi\,\rangle$, we must have:

$$1 = \|\psi\|^2 = \langle\, \psi \,|\, \psi \,\rangle = \int \mathrm{d}x\, |\psi(x)|^2 = \int \frac{\mathrm{d}p}{(2\pi\hbar)}\, |\tilde{\psi}(p)|^2\,. \tag{1.112}$$

A note on units here. $|\,\psi\,\rangle$ has no units. According to our conventions for the expansion of a general vector, the amplitude $\psi(x)$ has units of $1/\sqrt{L}$ and, since $\hbar$ has units of ML^2/T, the amplitude $\tilde{\psi}(p)$ has units of $\sqrt{L}$. This means that $|\,x\,\rangle$ as units of $1/\sqrt{L}$ and $|\,p\,\rangle$ has units of $\sqrt{L}$. The *operator* X has units of L and the *operator* P units of momentum (ML^2/T^2). These choices of units are just conventions; other choices work just as well.

1.5.1 Translation of the coordinate system

Suppose a coordinate frame Σ' is displaced, in one dimension, along the x-axis from a fixed frame Σ by an amount a. Then a point P is described by the coordinate x' in frame Σ' and by a point $x = x' + a$ in frame Σ. In this section, we want to find a unitary quantum operator which, when acting on the quantum coordinate operator X, will displace the quantum operator by an amount a. Using the commutation relations (1.109), we can construct such a unitary operator which does this displacement on quantum operators. We define $U(a)$ by:

$$U(a) = e^{-iPa/\hbar}\,, \tag{1.113}$$

where P is the momentum operator, then using Eq. (B.14) in appendix **??**, we find:

$$\begin{aligned} U^\dagger(a)\, X\, U(a) &= X + [\, iPa/\hbar, X\,] + \frac{1}{2}\,[\, iPa/\hbar, [\, iPa/\hbar, X\,]\,] + \cdots \\ &= X + a\,, \end{aligned} \tag{1.114}$$

and therefore we can write:

$$X\, U(a) = U(a)\,(\, X + a\,)\,.$$

So the operation of $X\,U(a)$ on the ket $|\,x'\,\rangle$ gives:

$$X\,\{\, U(a)\,|\,x'\,\rangle\,\} = U(a)\,(\, X + a\,)\,|\,x'\,\rangle = (\, x' + a\,)\,\{\, U(a)\,|\,x'\,\rangle\,\} = x\,\{\, U(a)\,|\,x'\,\rangle\,\}\,. \tag{1.115}$$

In other words, $U(a)\,|\,x'\,\rangle$ is an eigenvector of the operator X with eigenvalue x. That is:

$$U(a)\,|\,x'\,\rangle = |\,x\,\rangle\,, \qquad \text{or} \qquad |\,x'\,\rangle = U^\dagger(a)\,|\,x\,\rangle\,. \tag{1.116}$$

$U(a)$ is called a displacement operator. Then

$$\psi(x') = \langle\, x' \,|\, \psi \,\rangle = \langle\, x \,|\, U(a) \,|\, \psi \,\rangle = \langle\, x \,|\, \psi' \,\rangle = \psi'(x)\,, \tag{1.117}$$

where $|\,\psi'\,\rangle = U(a)\,|\,\psi\,\rangle$. So the function $\psi'(x)$ in the displaced coordinate system is defined by the relation:

$$\psi'(x) = \psi(x') = \psi(x - a)\,. \tag{1.118}$$

For infinitesimal displacements by an amount $a = \Delta x$, the displacement operator is given by the expansion:

$$U(\Delta x) = 1 - i\, P\Delta x/\hbar + \cdots\,. \tag{1.119}$$

Putting this into Eq. (1.117), we find:

$$\begin{aligned}\psi(x-\Delta x) &= \psi(x) - \frac{\partial\psi(x)}{\partial x}\,\Delta x + \cdots \\ &= \langle\, x\,|\,\big[\,1 - \frac{i}{\hbar}\,P\,\Delta x + \cdots\big]\,|\,\psi\,\rangle = \psi(x) - \frac{i}{\hbar}\,\langle\, x\,|\,P\,|\,\psi\,\rangle\,\Delta x + \cdots\,,\end{aligned} \tag{1.120}$$

so

$$\langle\, x\,|\,P\,|\,\psi\,\rangle = \frac{\hbar}{i}\,\frac{\partial\psi(x)}{\partial x}\,. \tag{1.121}$$

That is, the operator P acts as a derivative on a function in coordinate space.

We can also use these results to find the unitary connection between the $|\,x\,\rangle$ and $|\,p\,\rangle$ basis sets. We start by noting that Eq. (1.116) can be used to find $|\,x\,\rangle$ for any x, given the ket $|\,0\,\rangle$. That is, put $x' = 0$ so that $a = x$. This gives:

$$|\,x\,\rangle = U(x)\,|\,0\,\rangle = e^{-iPx/\hbar}\,|\,0\,\rangle\,. \tag{1.122}$$

so that:

$$\langle\, p\,|\,x\,\rangle = e^{-ipx/\hbar}\,\langle\, p\,|\,0\,\rangle = e^{-ipx/\hbar}\,. \tag{1.123}$$

Here we have set $\langle\, p\,|\,0\,\rangle = 1$. This is an arbitrary p-dependent normalization factor. This choice of normalization then gives:

$$\langle\, x\,|\,p\,\rangle = e^{ipx/\hbar}\,, \tag{1.124}$$

so that:

$$\psi(x) = \langle\, x\,|\,\psi\,\rangle = \int\frac{\mathrm{d}p}{(2\pi\hbar)}\,\langle\, x\,|\,p\,\rangle\,\langle\, p\,|\,\psi\,\rangle = \int\frac{\mathrm{d}p}{(2\pi\hbar)}\,e^{ipx/\hbar}\,\tilde{\psi}(p)\,, \tag{1.125}$$

which is just a Fourier transform of the function $\psi(x)$.

1.6 Measurement

The **state** of a system is described by a vector $|\,\psi\,\rangle$ in a vector space $\mathcal{V}$. Any vector that differs from $|\,\psi\,\rangle$ by a phase describes the same physical state. Observables are represented in quantum theory by Hermitian operators acting on vectors in $\mathcal{V}$. For example, observables for a single particle are the position, momentum, and spin. The values of these observable operators that can be measured are the eigenvalues and the probabilities of observing them for the state $|\,\psi\,\rangle$ are given by $\mathcal{P}_a = |\langle\, a\,|\,\psi\,\rangle|^2$. The probability of finding the particle with any value of the observables is called the **expectation value**, and is given by:

$$\langle\, A\,\rangle = \sum_a a\,|\langle\, a\,|\,\psi\,\rangle|^2 = \langle\,\psi\,|\,A\,|\,\psi\,\rangle\,.$$

The expectation value of A^2 is given by:

$$\langle\, A^2\,\rangle = \sum_a a^2\,|\langle\, a\,|\,\psi\,\rangle|^2 = \langle\,\psi\,|\,A^2\,|\,\psi\,\rangle\,,$$

with a similar relation for any power of the observable A. The mean uncertainty Δa in a measurement of A is given by:

$$(\Delta a)^2 = \sum_a (a - \langle\, A\,\rangle)^2\,|\langle\, a\,|\,\psi\,\rangle|^2 = \langle\,(A - \langle\, A\,\rangle)^2\,\rangle = \langle\, A^2\,\rangle - \langle\, A\,\rangle^2\,.$$

It is possible to set up devices to prepare the system to be in a particular state. For example, in a Stern-Gerlach type experiment, one can pass a beam, described by the state $|\,\psi\,\rangle$ through a device, which we call property A. The device then selects out beams with value a of the property A by blocking all other beams.

The experiment is described in quantum mechanics by the action of a projection operator $P_a = |\,a\,\rangle\langle\,a\,|$ on the initial state $|\,\psi\,\rangle$:

$$P_a\,|\,\psi\,\rangle = |\,a\,\rangle\,\langle\,a\,|\,\psi\,\rangle\,.$$

The act of blocking all other beams puts the system in a known state $|\,a\,\rangle$. For a second measurement of the property B, we have:

$$P_b P_a|\,\psi\,\rangle = |\,b\,\rangle\langle\,b\,|\,a\,\rangle\langle\,a\,|\,\psi\,\rangle\,,$$

and so on, for any number of measurements. If properties A and B do not commute, the measurement of A produces an eigenstate of A but the measurement of B puts the system into an eigenstate of B and erases the effect of the measurement of A. If properties A and B commute, then there is a common eigenvector $|\,ab\,\rangle$ and a common projection operator: $P_{ab} = |\,ab\,\rangle\langle\,ab\,|$ so that the experimental device can put the system into a eigenstate of both A and B.

Example 17. Let us take $A = X$ and $B = P$. Then since these two observables do not commute, then cannot be measured simultaneously. That is a measurement of first X and then P yields:

$$P_p\,P_x\,|\,\psi\,\rangle = |\,p\,\rangle\,\langle\,p\,|\,x\,\rangle\,\langle\,x\,|\,\psi\,\rangle = |\,x\,\rangle\,e^{-ipx/\hbar}\,\psi(x)\,, \tag{1.126}$$

On the other hand, a measurement of first P and then X yields:

$$P_x\,P_p\,|\,\psi\,\rangle = |\,x\,\rangle\,\langle\,x\,|\,p\,\rangle\,\langle\,p\,|\,\psi\,\rangle = |\,x\,\rangle\,e^{ipx/\hbar}\,\tilde{\psi}(p)\,, \tag{1.127}$$

an entirely different result than found in Eq. (1.126).

1.6.1 The uncertainty relation

Theorem 7 (Uncertainty principle)**.** *If A and B are two Hermitian operators and if $[\,A, B\,] = iC$, then the uncertainty of a common measurement of A and B for the state $|\,\psi\,\rangle$ is given by:*

$$(\Delta a)\,(\Delta b) \geq |\langle\,C\,\rangle|/2\,. \tag{1.128}$$

Proof. We first put:

$$\Delta A = A - \langle\,A\,\rangle\,, \qquad \Delta B = B - \langle\,B\,\rangle\,. \tag{1.129}$$

Then using the Schwartz inequality, we find:

$$\begin{aligned}(\Delta a)^2\,(\Delta b)^2 &= \langle\,(\Delta A)^2\,\rangle\,\langle\,(\Delta B)^2\,\rangle \\ &= \|\Delta A|\,\psi\,\rangle\|^2\,\|\Delta B|\,\psi\,\rangle\|^2 \geq |\langle\,\psi\,|\Delta A\,\Delta B|\,\psi\,\rangle|^2\,.\end{aligned} \tag{1.130}$$

Using

$$\begin{aligned}\Delta A\,\Delta B &= \frac{1}{2}[\,\Delta A\,\Delta B + \Delta B\,\Delta A\,] + \frac{i}{2}[\,\Delta A\,\Delta B - \Delta B\,\Delta A\,]/i \\ &= \frac{1}{2}[\,\Delta A\,\Delta B + \Delta B\,\Delta A\,] + \frac{i}{2}[\,A, B\,]/i \\ &= \frac{1}{2}\,[\,F + iC\,]\,.\end{aligned}$$

where F is given by:

$$F = \Delta A\,\Delta B + \Delta B\,\Delta A\,, \tag{1.131}$$

So from (1.130), we find:

$$\begin{aligned}\langle\,(\Delta A)^2\,\rangle\,\langle\,(\Delta B)^2\,\rangle = \frac{1}{4}[\langle\,F\,\rangle^2 + \langle\,C\,\rangle^2] \geq \frac{1}{4}\langle\,C\,\rangle^2\,,\\ \text{or,} \qquad \Delta a\,\Delta b \geq |\langle\,C\,\rangle|/2\,,\end{aligned}$$

which proves the theorem. □

Remark 10. A state of minimum uncertainty in the measurements is reached if the following holds:

- $\Delta B\,|\,\psi\,\rangle = \lambda\,\Delta A\,|\,\psi\,\rangle$, and
- $\langle\, F\,\rangle = 0$,

where λ is some constant. From the first requirement, we find that:

$$\langle\, \Delta A\,\Delta B\,\rangle = \lambda\,\langle\,\Delta A\,\rangle^2 = \lambda\,(\Delta a)^2$$
$$\langle\, \Delta B\,\Delta A\,\rangle = \langle\,\Delta B\,\rangle^2/\lambda = (\Delta b)^2/\lambda\,.$$

So adding and subtracting these last two equations gives:

$$\lambda\,(\Delta a)^2 + \frac{(\Delta b)^2}{\lambda} = \langle\, F\,\rangle = 0\,,$$
$$\lambda\,(\Delta a)^2 - \frac{(\Delta b)^2}{\lambda} = i\langle\, C\,\rangle\,.$$

Thus we find that

$$\lambda = i\frac{\langle\, C\,\rangle}{2(\Delta a)^2} = i\frac{\Delta b}{\Delta a}\,.$$

So the ket $|\,\psi\,\rangle$ which produces the minimum uncertainty in the product of the variances is given by the solution of:

$$\{\Delta b\,\Delta A + i\Delta a\,\Delta B\}|\,\psi\,\rangle = 0\,.$$

or

$$\{\,\Delta b\,A + i\Delta a\,B\,\}|\,\psi\,\rangle = \{\Delta b\,\langle\,A\,\rangle + i\Delta a\,\langle\,B\,\rangle\,\}|\,\psi\,\rangle\,. \tag{1.132}$$

That is, $|\,\psi\,\rangle$ is an eigenvector of the non-hermitian operator D, given by:

$$D = \Delta b\,A + i\Delta a\,B\,,$$

with complex eigenvalue d given by:

$$d = \Delta b\,\langle\,A\,\rangle + i\Delta a\,\langle\,B\,\rangle\,.$$

We have $D|\,\psi\,\rangle = d|\,\psi\,\rangle$. This state ψ is called a "coherent state" of the operators A and B.

Example 18. For the case when $A = X$ and $B = P$, with $[X, P] = i\hbar$. The minimum wave packet has $\Delta x\,\Delta p = \hbar/2$. Then Eq. (1.132) becomes:

$$\{\,\Delta p\,X + i\,\Delta x\,P\,\}\,|\,\psi\,\rangle = \{\,\Delta p\,\bar{x} + i\,\Delta x\,\bar{p}\,\}\,|\,\psi\,\rangle\,, \tag{1.133}$$

where $\bar{x} = \langle\,X\,\rangle$ and $\bar{p} = \langle\,P\,\rangle$. Operating on this equation on the left by $\langle\,x\,|$ gives the differential equation:

$$\left\{\,\Delta p\,x + \hbar\,\Delta x\,\frac{\mathrm{d}}{\mathrm{d}x}\,\right\}\psi(x) = \left\{\,\Delta p\,\bar{x} + i\,\Delta x\,\bar{p}\,\right\}\psi(x)\,,$$

which can be rearranged to give:

$$\left\{\frac{d}{dx} + \left[\frac{x-\bar{x}}{2\,(\Delta x)^2} - \frac{i\bar{p}}{\hbar}\right]\right\}\,\psi(x) = 0\,,$$

the solution of which is:

$$\psi(x) = \mathcal{N}\,\exp\left\{-\frac{(x-\bar{x})^2}{4\,(\Delta x)^2} + \frac{i}{\hbar}\,\bar{p}\,x\right\}\,, \tag{1.134}$$

where $\mathcal{N}$ is a normalization constant. Thus the wave function for the minimum uncertainty in position and momentum of the particle is a Gaussian wave packet.

1.7 Time in non-relativistic quantum mechanics

Time plays a special role in non-relativistic quantum mechanics. All vectors in the vector space which describe the physical system are functions of time. To put it another way, in non-relativistic quantum mechanics, time is considered to be a base manifold of one real dimension (t), and a different vector space $\mathcal{V}(t)$ is attached to this manifold at each value of the parameter t. Thus a vector describing the state of the system at time t is written as $|\,\Psi(t)\,\rangle$. A hermitian operator describing the property A for the vector space at time t is written as $A(t)$. Eigenvalues of this operator are written as a_i and eigenvectors of the operator are written as $|\,a_i, t\,\rangle$. Now one of Natures symmetries we want to preserve in a quantum theory is the inability to measure absolute time. By this, we mean that an observer with a clock measuring a time t does an experiment measuring property A with values a_i and probabilities $P_i(t) = |\,\langle\,\psi(t)\,|\,a_i, t\,\rangle\,|^2$, then an observer looking at the very same experiment but with a clock measuring a time $t' = t + \tau$ must measure exactly the same values a_i with the same probabilities $P_i(t') = |\,\langle\,\psi(t')\,|\,a_i, t'\,\rangle\,|^2$. Writing

$$|\,\psi(t')\,\rangle = U(\tau)\,|\,\psi(t)\,\rangle\,, \qquad \text{and} \qquad |\,a_i, t'\,\rangle = U(\tau)\,|\,a_i, t\,\rangle\,, \tag{1.135}$$

we see that $U(\tau)$ must be either linear and unitary or anti-linear and anti-unitary. Wigner proved this in the 1930's. For the case of time translations, the operator is linear and unitary. For infinitesimal time displacements $\Delta\tau$, the unitary operator representing this displacement is:

$$U(\Delta\tau) = 1 - \frac{i}{\hbar}\,H\,\Delta\tau + \cdots\,. \tag{1.136}$$

since $U(\Delta\tau)$ is unitary, we have introduced a factor of $i/\hbar$ so as to make H Hermitian with units of energy. From this point of view, $\hbar$ is a *necessary* factor. The negative sign is a convention. So from (1.135), we have:

$$\begin{aligned}\langle\,x\,|\,U(\Delta\tau)\,|\,\psi(t)\,\rangle &= \langle\,x\,|\,\Big\{\,1 - \frac{i}{\hbar}\,H\,\Delta\tau + \cdots\Big\}\,|\,\psi(t)\,\rangle \\ &= \langle\,x\,|\,\psi(t+\Delta\tau)\,\rangle = \psi(x,t) + \frac{\partial\psi(x,t)}{\partial t}\,\Delta\tau + \cdots\end{aligned} \tag{1.137}$$

where we have set $\psi(x,t) = \langle\,x\,|\,\Psi(t)\,\rangle$. So from (1.137), we find:

$$\langle\,x\,|\,H\,|\,\psi(t)\,\rangle = i\hbar\,\frac{\partial\psi(x,t)}{\partial t}\,. \tag{1.138}$$

If we choose H to be the total energy:

$$H = \frac{P^2}{2m} + V(X)\,, \tag{1.139}$$

then (1.138) becomes:

$$\Big\{-\frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial x^2} + V(x)\Big\}\,\psi(x,t) = i\hbar\,\frac{\partial\psi(x,t)}{\partial t}\,, \tag{1.140}$$

which is called Schrödinger's equation. We will re-derive these results from different points of view in the following chapters.

References

[1] B. D. Serot, "Introduction to Quantum Mchanics," (May, 1997).

[2] M. L. Boaz, *Mathematical Methods in the Physical Sciences* (John Wiley & Sons, New York, NY, 1983).

Chapter 2

Canonical quantization

We show in this chapter how to construct quantum theories from classical systems using canonical quantization postulates, which can be formulated for most physical systems of interest. The canonical formulation is stated in the form of generalized coordinates and an action, from which equations of motion are obtained. There is a relationship between cannonical transformations in classical physics and unitary transformations in quantum mechanics, so that exactly which canonical variables are used in the quantization procedure are irrelevant and give the same experimental results. So it is sometimes useful to try to find "good" classical variables first before quantization of the system is carried out. By good, we mean variables that describe the system in a simple way. Since in this method classical variables are replaced by non-commuting quantum operators, there can be ordering ambiguities in carrying out a canonical quantization procedure. Quantum mechanics does not tell us how to resolve such ambiguities, and so one must be resigned to realize that quantum systems can be reduced to classical systems, but the opposite may not be true. Only experiment will tell us what is the correct quantum realization of a system. In addition, some systems have *no* classical analog at all! For example a Fermi oscillator has no classical description. In general quantum systems containing *anti*-commuting operators belong to this class. Nevertheless, we shall see by some examples, that canonical quantization is very often a useful tool to obtain the correct quantum mechanics, and so we will study this method in this chapter.

2.1 Classical mechanics review

We start by considering a classical system, described by a Lagrangian which is a function of the generalized coordinates $q_i(t)$, velocities, $\dot{q}_i(t)$, for $i = 1, \ldots, n$, and time. The classical action is given by:

$$S[q] = \int L(q, \dot{q}, t)\, dt\,, \tag{2.1}$$

and is a functional of the paths of the system $q_i(t)$ in q-space. We show in Section 2.1.1 below that Lagrange's equations of motion are obtained by requiring the action to be stationary under variation of the functional form of the paths $q_i(t)$ in q-space, with no variation of the end points:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0\,, \qquad \text{for } i = 1, \ldots, n.$$

The canonical momentum p_i, conjugate to q_i, is given by:

$$p_i = \frac{\partial L}{\partial \dot{q}_i}\,, \qquad \text{for } i = 1, \ldots, n.$$

The Hamiltonian is defined by the transformation,

$$H(q,p,t) = \sum_{i=1}^{n} p_i \dot{q}_i - L(q,\dot{q},t) \,,$$

and Hamilton's equations of motion are:

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} = \{\, p_i, H \,\} \,, \qquad \dot{q}_i = +\frac{\partial H}{\partial p_i} = \{\, q_i, H \,\} \,,$$

for $i = 1, \ldots, n$. These equations are equivalent to Newton's laws. Here the curly brackets are classical Poisson brackets, not to be confused with quantum mechanical anti-commutators, and are defined by:

$$\{\, A, B \,\} = \sum_{i=1}^{n} \left(\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial B}{\partial q_i}\frac{\partial A}{\partial p_i} \right) \,.$$

In particular, we have:

$$\{\, q_i, q_j \,\} = 0 \,, \qquad \{\, p_i, p_j \,\} = 0 \,, \qquad \{\, q_i, p_j \,\} = \delta_{ij} \,.$$

For any function F of p, q, and t, we have:

$$\begin{aligned} \frac{dF(p,q,t)}{dt} &= \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q_i}\dot{q}_i + \frac{\partial F}{\partial p_i}\dot{p}_i \right) + \frac{\partial F}{\partial t} \,, \\ &= \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial q_i}\frac{\partial F}{\partial p_i} \right) + \frac{\partial F}{\partial t} \,, \\ &= \{\, F, H \,\} + \frac{\partial F}{\partial t} \,. \end{aligned} \tag{2.2}$$

Constants of the motion are those for which $\mathrm{d}F(q,p,t)/\mathrm{d}t = 0$. In particular, for $F = H$, the Hamiltonian, we find:

$$\frac{dH(q,p,t)}{dt} = \frac{\partial H(q,p,t)}{\partial t} \,. \tag{2.3}$$

Thus the Hamiltonian is a constant of the motion if H doesn't dependent *explicitly* on time.

2.1.1 Symmetries of the action

We study in this section the consequences of classical symmetries of the action. We suppose that the action is of the form given in Eq. (2.1). We consider infinitesimal variations of time and the generalized coordinates of the form:

$$t' = t + \delta t(t) \,, \qquad q_i'(t') = q_i(t) + \Delta q_i(t) \,, \tag{2.4}$$

where, to first order,

$$\Delta q_i(t) = \delta q_i(t) + \dot{q}_i(t)\,\delta t(t) \,, \qquad \delta q_i(t) = q_i'(t) - q_i(t) \,. \tag{2.5}$$

Here $\Delta q_i(t)$ is the *total* change in $q_i(t)$ whereas $\delta q_i(t)$ is a change in *functional form*. To first order in δt, that the differential time element $\mathrm{d}t$ changes by:

$$\mathrm{d}t' = (\, 1 + \dot{\delta} t(t) \,)\, \mathrm{d}t \,. \tag{2.6}$$

The change in the action under this variation is given by:

$$\begin{aligned}
\Delta S[q] &= \int L(q',\dot{q}',t')\,\mathrm{d}t' - \int L(q,\dot{q},t)\,\mathrm{d}t \\
&= \int\Big\{\frac{\partial L}{\partial q_i}\,\Delta q_i(t) + \frac{\partial L}{\partial \dot{q}_i}\,\Delta\dot{q}_i(t) + \frac{\partial L}{\partial t}\,\delta t(t) + L\,\dot{\delta t}(t)\Big\}\,\mathrm{d}t \\
&= \int\Big\{\frac{\partial L}{\partial q_i}\,\delta q_i(t) + \frac{\partial L}{\partial \dot{q}_i}\,\delta\dot{q}_i(t) + \Big[\frac{\partial L}{\partial q_i}\,\dot{q}_i(t) + \frac{\partial L}{\partial \dot{q}_i}\,\ddot{q}_i(t) + \frac{\partial L}{\partial t}\Big]\,\delta t(t) + L\,\dot{\delta t}(t)\Big\}\,\mathrm{d}t \\
&= \int\Big\{\frac{\partial L}{\partial q_i}\,\delta q_i(t) + \frac{\partial L}{\partial \dot{q}_i}\,\delta\dot{q}_i(t) + \frac{\mathrm{d}L}{\mathrm{d}t}\,\delta t(t) + L\,\dot{\delta t}(t)\Big\}\,\mathrm{d}t \\
&= \int\Big\{\frac{\partial L}{\partial q_i}\,\delta q_i(t) + \frac{\partial L}{\partial \dot{q}_i}\,\frac{\mathrm{d}\,\delta q_i(t)}{\mathrm{d}t} + \frac{\mathrm{d}}{\mathrm{d}t}\,[\,L\,\delta t(t)\,]\Big\}\,\mathrm{d}t \\
&= \int\Big\{\Big[\frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big)\Big]\,\delta q_i(t) + \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial L}{\partial \dot{q}_i}\,\delta q_i(t) + L\,\delta t(t)\Big]\Big\}\,\mathrm{d}t \\
&= \int\Big\{\Big[\frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big)\Big]\,\delta q_i(t) + \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial L}{\partial \dot{q}_i}\,\Delta q_i(t) - \Big(\frac{\partial L}{\partial \dot{q}_i}\,\dot{q}_i - L\Big)\,\delta t(t)\Big]\Big\}\,\mathrm{d}t \\
&= \int\Big\{\Big[\frac{\partial L}{\partial q_i} - \dot{p}_i\Big]\,\delta q_i(t) + \frac{\mathrm{d}}{\mathrm{d}t}\Big[p_i\,\Delta q_i(t) - H\,\delta t(t)\Big]\Big\}\,\mathrm{d}t
\end{aligned} \tag{2.7}$$

In the last line, we have set $p_i = \partial L/\partial\dot{q}_i$ and $H = p_i\dot{q}_i - L$. So if we require the action to be stationary with respect to changes $\delta q_i(t)$ in the functional form of the paths in q-space, with no variations at the end points so that $\delta q_i(t_1) = \delta q_i(t_2) = 0$ and no changes in the time variable $\delta t(t) = 0$ so that $\Delta q_i(t) = \delta q_i(t)$, then the second term above vanishes, and we find Lagrange's equations of motion for the $q_i(t)$ variables:

$$\dot{p}_i = \frac{\partial L}{\partial q_i}\,, \qquad \text{for } i = 1,\ldots,n. \tag{2.8}$$

On the other hand, if the $q_i(t)$ variables satisfy Lagrange's equation, then the first term vanishes, and if the action is invariant under the variations $\Delta q_i(t)$ and $\delta t(t)$, then the second term requires that

$$p_i\,\Delta q_i(t) - H\,\delta t(t)\,, \tag{2.9}$$

are constants of the motion. What we have shown here is that symmetries of the action lead to conservation laws for the generators of the transformation.

Example 19. If the action is invariant under time translations, then $\delta t(t) = \delta\tau$ and $\Delta q_i(t) = 0$ for all i. Then Eq. (2.9) shows that the Hamiltonian H is a constant of the motion. This is in agreement with our statement that H is conserved if the Lagrangian does not depend *explicitly* on time.

If the action is invariant under space translation of all coordinates q_i, then $\delta t(t) = 0$ and $\Delta q_i(t) = \delta a$ for all i. Then Eq. (2.9) shows that the total momentum of the system is conserved, and that

$$P = \sum_i p_i\,, \tag{2.10}$$

is a constant of the motion.

2.1.2 Galilean transformations

Let us now specialize to a system of N particles of mass m described by $n = 3N$ generalized cartesian coordinates: $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$, and with interactions between the particles that depend only on the magnitude of the distance between them. The Lagrangian for this system is given by:

$$L(\mathbf{x},\dot{\mathbf{x}}) = \frac{1}{2}\sum_{i=1}^{N} m\,|\,\dot{\mathbf{x}}_i\,|^2 - \frac{1}{2}\sum_{i,j=1(j\neq i)}^{N} V(|\mathbf{x}_i - \mathbf{x}_j|)\,, \tag{2.11}$$

The canonical momentum is given by: $\mathbf{p}_i = m\,\dot{\mathbf{x}}_i$, and the equations of motion are:

$$\dot{\mathbf{p}}_i = -\boldsymbol{\nabla}_i \sum_{j=1(j\neq i)}^{N} V(|\mathbf{x}_i - \mathbf{x}_j|)\,, \qquad \text{for } i = 1, \ldots, N. \tag{2.12}$$

If $V(\mathbf{x})$ depends only on the difference of coordinates of pairs of particles, the action for this Lagrangian is stationary with respect to infinitesimal Galilean transformations of the form:

$$\Delta\mathbf{x}_i(t) = (\,\hat{\mathbf{n}} \times \mathbf{x}_i\,)\Delta\theta + \Delta\mathbf{v}\,t + \Delta\mathbf{a}\,, \qquad \delta t(t) = \Delta\tau\,, \tag{2.13}$$

for all $i = 1, \ldots, N$.

Exercise 1. Prove that the action for the many-particle Lagrangian (2.11) is invariant under Galiliean transformations.

From (2.9) and (2.13), the conserved generators are then:

$$\sum_{i=1}^{N} \mathbf{p}_i \cdot \Delta\mathbf{x}_i(t) - H\,\delta t(t) = -\Delta\theta\,\hat{\mathbf{n}} \cdot \mathbf{J} - \Delta\mathbf{v} \cdot \mathbf{K} + \Delta\mathbf{a} \cdot \mathbf{P} - \Delta\tau\,H\,, \tag{2.14}$$

where

$$\mathbf{J} = \sum_{i=1}^{N} \mathbf{x}_i \times \mathbf{p}_i\,, \qquad \mathbf{K} = \sum_{i=1}^{N} t\,\mathbf{p}_i\,, \tag{2.15}$$

$$\mathbf{P} = \sum_{i=1}^{N} \mathbf{p}_i\,, \qquad H = \sum_{i=1}^{N} \mathbf{p}_i \cdot \dot{\mathbf{x}}_i - L\,. \tag{2.16}$$

So the set of ten classical generators $(\mathbf{J}, \mathbf{K}, \mathbf{P}, H)$ are all conserved if the action is invariant under Galilean transformations.

2.2 Canonical quantization postulates

The canonical quantization method attempts to create a quantum system from the classical description in terms of generalized coordinated, by associating the classical generalized coordinates and momenta to Hermitian operators in a linear vector space. The associated operators are considered to be observables of the system. These observable operators obey a commutation algebra. Possible *states* of the system are described by vectors in this space. The dynamics of the system are found by mapping Poisson bracket relations in the classical system to commutation relations in the quantum system. This mapping is best described in what is called the Heisenberg picture, and is what we discuss in Section 2.2.1 below. A second way of looking at the dynamics is called the Schrödinger picture, and is discussed Section 2.2.2.

In the Heisenberg picture, the observable operators change with time, moving in relation to the basis vectors in the vector space. The state of the system, on the other hand, remains fixed. In the Schrödinger picture, the observable operators remain fixed in space, but the state of the system changes with time. Both pictures are equivalent and can be made to coincide at $t = 0$.

For the remainder of this chapter, we will assume that the Hamiltonian does not dependent explicitly on time.

2.2.1 The Heisenberg picture

The canonical quantization postulates are easily stated in the Heisenberg picture using the Hamiltonian formalism from classical mechanics. These postulates are:

- The generalized coordinates $q_i(t)$ and canonical momenta $p_i(t)$ map to hermitian operators in quantum mechanics:
$$q_i(t) \mapsto Q_i(t)\,, \qquad p_i(t) \mapsto P_i(t)\,, \tag{2.17}$$
- and the classical Poisson brackets map to commutators of operators in quantum mechanics, divided by $i\hbar$:
$$\{\, a(q,p,t), b(q,p,t)\,\} \mapsto \frac{[\,A(Q,P,t), B(Q,P,t)\,]}{i\hbar}\,. \tag{2.18}$$

In particular, at any time t, the operators $Q_i(t)$ and $P_i(t)$ obey the equal time commutation relations:

$$[\,Q_i(t), Q_j(t)\,] = 0\,, \qquad [\,P_i(t), P_j(t)\,] = 0\,, \qquad [\,Q_i(t), P_j(t)\,] = i\hbar\,\delta_{ij}\,. \tag{2.19}$$

The equations of motion in the Heisenberg representation are then described by the operator equations:

$$\begin{aligned} i\hbar\frac{\mathrm{d}Q_i(t)}{\mathrm{d}t} &= [\,Q_i(t), H(Q(t),P(t),t)\,]\,,\\ i\hbar\frac{\mathrm{d}P_i(t)}{\mathrm{d}t} &= [\,P_i(t), H(Q(t),P(t),t)\,]\,. \end{aligned} \tag{2.20}$$

Eqs. (2.20) were called by Dirac the Heisenberg equations of motion. Thus at any time t, we can simultaniously diagonalize all the $Q_i(t)$ operators and (separately) all the $P_i(t)$ operators:

$$\begin{aligned} Q_i(t)\,|\,q,t\,\rangle &= q_i\,|\,q,t\,\rangle\,,\\ P_i(t)\,|\,p,t\,\rangle &= p_i\,|\,p,t\,\rangle\,, \end{aligned} \tag{2.21}$$

where $|\,q,t\,\rangle$ and $|\,p,t\,\rangle$ stand for the set:

$$\begin{aligned} |\,q,t\,\rangle &= |\,q_1,q_2,\ldots,q_n,t\,\rangle\,,\\ |\,p,t\,\rangle &= |\,p_1,p_2,\ldots,p_n,t\,\rangle\,. \end{aligned} \tag{2.22}$$

Here q_i and p_i are real, with ranges and normalizations decided by the physical situation. Note that the eigen*vectors* of the operators $Q_i(t)$ and $P_i(t)$ depend on time, but the eigen*values* do not. The eigenvectors have the same spectrum for all time.

The construction of the the quantum mechanical Hamiltonian operator $H(Q(t),P(t),t)$ from the classical Hamiltonian is usually straightforward and leads to a Hermitian operator in most cases. However, there *can* be ordering problems involving non-commuting operators, such as $Q(t)$ and $P(t)$, in which case some method must be used to make the Hamiltonian Hermitian. We must require H to be Hermitian in order to conserve probability.

It is easy to show that the solution of the Heisenberg equations of motion, Eqs. (2.20), is given by:

$$Q_i(t) = U^\dagger(t)\,Q_i\,U(t)\,, \qquad \text{and} \qquad P_i(t) = U^\dagger(t)\,P_i\,U(t)\,, \tag{2.23}$$

where we have set $Q_i \equiv Q_i(0)$ and $P_i \equiv P_i(0)$ and where $U(t)$, the time-development operator, is the solution of the equation:

$$i\hbar\,\frac{\partial U(t)}{\partial t} = U(t)\,H(Q(t),P(t),t)\,, \qquad \text{and} \qquad -\,i\hbar\,\frac{\partial U^\dagger(t)}{\partial t} = H(Q(t),P(t),t)\,U^\dagger(t)\,, \tag{2.24}$$

with $U(0) = 1$. But since

$$H(Q(t),P(t),t) = U^\dagger(t)\,H(Q,P,t)\,U(t)\,, \tag{2.25}$$

so that $U(t)\,H(Q(t),P(t),t) = H(Q,P,t)\,U(t)$, we can write Eqs. (2.24) as:

$$i\hbar\,\frac{\partial U(t)}{\partial t} = H(Q,P,t)\,U(t)\,, \qquad \text{and} \qquad -\,i\hbar\,\frac{\partial U^\dagger(t)}{\partial t} = U^\dagger(t)\,H(Q,P,t)\,. \tag{2.26}$$

Now since $H(Q,P,t)$ is Hermitian, the time development operators are unitary:

$$\frac{d}{dt}\{U^\dagger(t)U(t)\} = 0\,, \qquad \Rightarrow \qquad U^\dagger(t)U(t) = 1\,. \tag{2.27}$$

That is probability is conserved even if energy is not. If the Hamiltonian is independent of time explicitly, the time-development operator $U(t)$ has a simple solution:

$$U(t) = e^{-iH(Q,P)t/\hbar}\,, \qquad U^\dagger(t) = e^{+iH(Q,P)t/\hbar}\,. \tag{2.28}$$

We discuss the case when the Hamiltonian has an explicit time dependence in the next chapter in Section 4.2.

From the eigenvalue equations (2.21), we find:

$$\begin{aligned} U^\dagger(t)\,Q_i\,U(t)\,|\,q,t\,\rangle &= q_i\,|\,q,t\,\rangle\,,\\ U^\dagger(t)\,P_i\,U(t)\,|\,p,t\,\rangle &= p_i\,|\,p,t\,\rangle\,. \end{aligned} \tag{2.29}$$

Operating on these equations on the left by $U(t)$ gives:

$$\begin{aligned} Q_i\,\big\{\,U(t)\,|\,q,t\,\rangle\,\big\} &= q_i\,\big\{\,U(t)\,|\,q,t\,\rangle\,\big\}\,,\\ P_i\,\big\{\,U(t)\,|\,p,t\,\rangle\,\big\} &= p_i\,\big\{\,U(t)\,|\,p,t\,\rangle\,\big\}\,, \end{aligned} \tag{2.30}$$

which means that:

$$\begin{aligned} U(t)\,|\,q,t\,\rangle &= |\,q,0\,\rangle \equiv |\,q\,\rangle\,,\\ U(t)\,|\,p,t\,\rangle &= |\,p,0\,\rangle \equiv |\,p\,\rangle\,, \end{aligned} \tag{2.31}$$

or

$$\begin{aligned} |\,q,t\,\rangle &= U^\dagger(t)\,|\,q\,\rangle\,,\\ |\,p,t\,\rangle &= U^\dagger(t)\,|\,p\,\rangle\,, \end{aligned} \tag{2.32}$$

from which we find:

$$\begin{aligned} H(Q,P,t)\,|\,q,t\,\rangle &= -i\hbar\,\frac{\partial}{\partial t}\,|\,q,t\,\rangle\,,\\ H(Q,P,t)\,|\,p,t\,\rangle &= -i\hbar\,\frac{\partial}{\partial t}\,|\,p,t\,\rangle\,. \end{aligned} \tag{2.33}$$

So in the Heisenberg picture, the base vectors change in time according to the unitary operator $U^\dagger(t)$.

Any operator function of $Q(t)$, $P(t)$ and t in the Heisenberg representation can be written as:

$$F(\,Q(t),P(t),t\,) = F(\,U^\dagger(t)\,Q\,U(t),\,U^\dagger(t)\,P\,U(t),\,t\,) = U^\dagger(t)\,F(\,Q,P,t\,)\,U(t)\,.$$

Then, using Eqs. (2.24), the total time derivative of the operator $F(\,Q(t),P(t),t\,)$ is given by:

$$\begin{aligned} \frac{\mathrm{d}F(\,Q(t),P(t),t\,)}{\mathrm{d}t} &= \frac{\partial U^\dagger(t)}{\partial t}\,F(\,Q,P,t\,)\,U(t) + U^\dagger(t)\,\frac{\partial F(\,Q,P,t\,)}{\partial t}\,U(t) + U^\dagger(t)\,F(\,Q,P,t\,)\,\frac{\partial U(t)}{\partial t}\\ &= \big\{\,U^\dagger(t)\,F(\,Q,P,t\,)\,U(t)\,H(Q(t),P(t),t) - H(Q(t),P(t),t)\,U^\dagger(t)\,F(\,Q,P,t\,)\,U(t)\,\big\}/(i\hbar)\\ &\quad + U^\dagger(t)\,\frac{\partial F(\,Q,P,t\,)}{\partial t}\,U(t)\\ &= \frac{[\,F(\,Q(t),P(t),t\,),H(Q(t),P(t),t)\,]}{i\hbar} + \frac{\partial F(\,Q(t),P(t),t\,)}{\partial t}\,, \end{aligned} \tag{2.34}$$

in agreement with the classical result, Eq. (2.2), with the arguments replaced by time-dependent operators and the Poisson Bracket replaced by a commutator divided by $i\hbar$. The last partial derivative term in (2.34)

means to calculate the partial derivative of the *explicit* dependence of $F(Q(t),P(t),t)$ with respect to time. If we set $F(Q(t),P(t),t) = H(Q(t),P(t))$, we find that if the Hamiltonian is independent explicitly of time,

$$\frac{\mathrm{d}H(Q(t),P(t))}{\mathrm{d}t} = \frac{\partial H(Q(t),P(t))}{\partial t} = 0\,, \tag{2.35}$$

and is conserved. That is $H(Q(t),P(t)) = H(Q,P)$ for all t. In this section, we have tried to be very careful with the time-dependent arguments of the operators, and to distinguish between operators like $Q_i(t)$ and Q_i, the last of which is time-independent. Eqs. (2.24) and Eqs. (2.26) are examples of the importance of making this distinction.

In the Heisenberg picture, the goal is to solve the equation of motion of the operators using Eqs. (2.20) with initial values of the operators. This gives a complete description of the system, but very often, this is a difficult job, and one resorts to finding equations of motion for average values of the operators and their moments. However we will see another way to solve for the dynamics in the next section.

2.2.2 The Schrödinger picture

The probability amplitude of finding the system in the state ψ with coordinates q at time t is:

$$\psi(q,t) = \langle\, q,t\,|\,\psi\,\rangle = \langle\, q\,|\,U(t)\,|\,\psi\,\rangle = \langle\, q\,|\,\psi(t)\,\rangle\,, \tag{2.36}$$

where we have set:

$$|\,\psi(t)\,\rangle = U(t)\,|\,\psi\,\rangle\,. \tag{2.37}$$

Differentiating both sides of this equation with respect to t, and using Eq. (2.26), gives Schrödinger's equation:

$$H(Q,P,t)\,|\,\psi(t)\,\rangle = i\hbar\,\frac{\partial}{\partial t}\,|\,\psi(t)\,\rangle\,. \tag{2.38}$$

Here Q and P *have no time-dependence*. In the Schrödinger picture, the state vector $|\,\psi(t)\,\rangle$ moves but the operators Q and P and the base vectors remain stationary. The effort then is to solve Schrödinger's equation in this picture. The length of the state vector is conserved:

$$\langle\,\psi(t)\,|\,\psi(t)\,\rangle = \langle\,\psi\,|\,U^\dagger(t)\,U(t)\,|\,\psi\,\rangle = \langle\,\psi\,|\,\psi\,\rangle\,, \tag{2.39}$$

so that the probability of finding the system in *some* state is always unity. Note that Eq. (2.33) looks like Schrödinger's equation (2.38) but for a negative sign on the right-hand side. That is, it evolves "backward" in time. Wave functions in coordinate and momentum space are defined by:

$$\begin{aligned}\psi(q,t) &= \langle\, q\,|U(t)|\,\psi\,\rangle = \langle\, q\,|\,\psi(t)\,\rangle = \langle\, q,t\,|\,\psi\,\rangle\,,\\ \tilde\psi(p,t) &= \langle\, p\,|U(t)|\,\psi\,\rangle = \langle\, p\,|\,\psi(t)\,\rangle = \langle\, p,t\,|\,\psi\,\rangle\,.\end{aligned}$$

For a Hamiltonian of the form:

$$H(Q,P) = \frac{P^2}{2m} + V(Q)\,, \tag{2.40}$$

the coordinate representation wave function satisfies the differential equation:

$$\left\{-\frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial q^2} + V(q)\right\}\psi(q,t) = i\hbar\,\frac{\partial\psi(q,t)}{\partial t}\,. \tag{2.41}$$

When the Hamiltonian is independent explicitly of time, the solution for the state vector can be found in a simple way in a representation of the eigenvectors of the Hamiltonian operator. For example, let $|\,E_i\,\rangle$ be an eigenvector of $H(Q,P)$ with eigenvalue E_i:

$$H(Q,P)\,|\,E_i\,\rangle = E_i\,|\,E_i\,\rangle\,, \tag{2.42}$$

Then we find:

$$\psi(E_i,t) = \langle\, E_i \,|\, \psi(t) \,\rangle = \langle\, E_i \,|\, e^{-iHt/\hbar} \,|\, \psi \,\rangle = e^{-iE_it/\hbar} \langle\, E_i \,|\, \psi \,\rangle = e^{-iE_it/\hbar}\, \psi(E_i)\,, \tag{2.43}$$

where $\psi(E_i)$ is the value of $\psi(E_i,t)$ at $t=0$. So the state vector in the energy representation just oscillates with frequency $e^{-iE_it/\hbar}$ for each energy mode. The vector itself is given by a sum over all energy eigenstates:

$$|\, \psi(t) \,\rangle = \sum_i \psi(E_i)\, e^{-iE_it/\hbar} \,|\, E_i \,\rangle\,. \tag{2.44}$$

The average value of operators can be found in both the Heisenberg and Schrödinger picture:

$$\langle\, F(t) \,\rangle_\psi = \langle\, \psi \,|\, F(\, Q(t), P(t), t\,) \,|\, \psi \,\rangle = \langle\, \psi(t) \,|\, F(Q,P,t) \,|\, \psi(t) \,\rangle\,.$$

In the Schrödinger picture, we solve the time-dependent Schrödinger's equation for $|\, \psi(t) \,\rangle$. In the coordinate representation, this amounts to solving a partial differential equation. Since we know a lot about solutions of partial differential equations, this method is sometimes easier to use than the Heisenberg picture. The quantum dynamics of a single particle in three-dimensions is usually studied in the Schrödinger picture whereas the dynamics of multiple particles are usually studied in the Heisenberg picture using quantum field theory.

We now turn to some simple examples.

Example 20. Free particle: The Hamiltonian for a free particle with mass m in one-dimension is:

$$H(Q,P) = \frac{P^2}{2m}\,, \qquad \text{and} \qquad [\, Q(t), P(t) \,] = i\hbar\,. \tag{2.45}$$

(i) We first solve this problem in the Heisenberg picture. The Heisenberg equations of motion give:

$$\begin{aligned} \dot{Q}(t) &= [\, Q(t), H \,]/(i\hbar) = P/(2m)\,, \\ \dot{P}(t) &= [\, P(t), H \,]/(i\hbar) = 0\,, \end{aligned} \tag{2.46}$$

which have the solutions:

$$P(t) = P_0\,, \qquad \text{and} \qquad Q(t) = Q_0 + \frac{P_0}{m}\, t\,. \tag{2.47}$$

So the average values of position and momentum are given by:

$$\begin{aligned} \langle\, P(t) \,\rangle &= \langle\, P_0 \,\rangle = p_0\,, \\ \langle\, Q(t) \,\rangle &= \langle\, Q_0 \,\rangle + \frac{\langle\, P_0 \,\rangle}{m}\, t = q_0 + \frac{p_0}{m}\, t\,, \end{aligned} \tag{2.48}$$

as expected from the classical result. Now since

$$[\, Q(t), Q_0 \,] = [\, Q_0 + P_0\, \frac{t}{m}, Q_0 \,] = -i\hbar \frac{t}{m}\,, \tag{2.49}$$

we find from the minimum uncertainty principle (Theorem 7 on page 25) that

$$\Delta q(t)\, \Delta q_0 \geq \frac{\hbar\, t}{2m}\,, \qquad \text{or} \qquad \Delta q(t) \geq \frac{\hbar\, t}{2m\, \Delta q_0}\,. \tag{2.50}$$

So the uncertainty of the position of the particle grows with time for *any* initial state. The uncertainties in momentum and position must be calculated from the relation:

$$\begin{aligned} (\, \Delta p(t)\,)^2 &= \langle\, (\, P(t) - p_0\,)^2 \,\rangle = \langle\, P^2(t) \,\rangle - p_0^2 = \langle\, P_0^2 \,\rangle - p_0^2 = (\, \Delta p_0\,)^2\,, \\ (\, \Delta q(t)\,)^2 &= \langle\, (\, Q(t) - q_0\,)^2 \,\rangle = \langle\, Q^2(t) \,\rangle - q_0^2 = \langle\, Q_0^2 + (Q_0P_0 + P_0Q_0)\frac{t}{m} + P_0^2\, \frac{t^2}{m^2} \,\rangle - q_0^2 \\ &= (\, \Delta q_0\,)^2 + C\, \frac{t}{m} + \{\, (\Delta p_0)^2 + p_0^2 \,\}\, \frac{t^2}{m^2}\,, \end{aligned} \tag{2.51}$$

where $C = \langle\, Q_0P_0 + P_0Q_0 \,\rangle$. For a coherent state $C = 2p_0q_0$, so for a coherent state:

$$(\,\Delta q(t)\,)^2 = (\,\Delta q_0\,)^2 - q_0^2 + \Big(\, q_0 + p_0\,\frac{t}{m}\,\Big)^2 + \Big(\,\Delta p_0\,\frac{t}{m}\,\Big)^2 \geq \Big[\,\frac{\hbar\, t}{2m\,\Delta q_0}\,\Big]^2\,, \tag{2.52}$$

in agreement with (2.50). So the uncertainty in momentum does not change with time, but the uncertainty in position always grows.

(ii) In the Schrödinger picture, we want to solve Schrödinger's equation, given by:

$$\frac{P^2}{2m}\,|\,\psi(t)\,\rangle = i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\,\psi(t)\,\rangle\,.$$

It is simpler if we solve this equation in the momentum representation. Multiplying through on the left by $\langle\, p\,|$, we get:

$$\frac{p^2}{2m}\,\tilde\psi(p,t) = i\hbar\frac{\partial\tilde\psi(p,t)}{\partial t}\,, \tag{2.53}$$

where $\tilde\psi(p,t) = \langle\, p\,|\,\psi(t)\,\rangle$. The solution of (2.53) is:

$$\tilde\psi(p,t) = \tilde\psi(p,0)\,\exp\left\{-\frac{i}{\hbar}\frac{p^2}{2m}\,t\right\}\,. \tag{2.54}$$

So the average value of the position is given by:

$$\begin{aligned}\langle\, Q(t)\,\rangle &= \frac{-i}{\hbar}\int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi\hbar}\,\tilde\psi^*(p,t)\frac{\partial\tilde\psi(p,t)}{\partial p}\\ &= \int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi\hbar}\left\{\,\tilde\psi^*(p,0)\left(\,\frac{-i}{\hbar}\,\frac{\partial}{\partial p}\,\tilde\psi(p,0)\right) + \frac{p\,t}{m}\,|\,\tilde\psi(p,0)\,|^2\,\right\}\\ &= q_0 + \frac{p_0}{m}\,t\,,\end{aligned} \tag{2.55}$$

in agreement with the result (2.48) in the Heisenberg picture. We will not bother to calculate the uncertainties in position and momentum in the Schrödinger equation.

2.3 Canonical transformations

The quantization rules we have used apply to any physical system described by canonical coordinates (with some restrictions to be described below). However, we know that we can transform the classical system by very general transformations to new coordinates which preserve Poisson brackets relations between the transformed coordinates and momenta and the form of Hamilton's equations. These canonical transformations to new coordiantes provide a completely equivalent description of the classical system. Thus, we must be able to quantize the system in any canonically equivalent system of coordinates, and obtain the same physics. What we need to show is that for every classical canonical transformation, we can find a unitary transformation in quantum mechanics which effects the change of coordinates. That is, if the classical transformation is given by:

$$\begin{aligned}q_i' &= q_i'(q,p,t)\,,\\ p_i' &= p_i'(q,p,t)\,,\end{aligned} \tag{2.56}$$

and is invertable, with (q,p) satisfying Hamilton's equations,

$$\begin{aligned}\dot q_i &= +\frac{\partial H(q,p,t)}{\partial p_i} = \{q_i, H(q,p,t)\}_{(q,p)}\,,\\ \dot p_i &= -\frac{\partial H(q,p,t)}{\partial q_i} = \{p_i, H(q,p,t)\}_{(q,p)}\end{aligned}$$

for some Hamiltonian $H(q,p,t)$, then the transformation (2.56) is canonical if we can find some new Hamiltonian $H'(q',p',t)$ such that the new set of coordinates and momentum (q',p') satisfy:

$$\begin{aligned}\dot{q}'_i &= \frac{\partial q'_i}{\partial q_j}\dot{q}_j + \frac{\partial q'_i}{\partial p_j}\dot{p}_j + \frac{\partial q'_i}{\partial t} = \{q'_i, H(q,p,t)\}_{(q,p)} + \frac{\partial q'_i}{\partial t} \\ &= +\frac{\partial H'(q',p',t)}{\partial p'_i} \qquad = \{q'_i, H'(q',p',t)\}_{(q',p')}\,, \\ \dot{p}'_i &= \frac{\partial p'_i}{\partial q_j}\dot{q}_j + \frac{\partial p'_i}{\partial p_j}\dot{p}_j + \frac{\partial p'_i}{\partial t} = \{p'_i, H(q,p,t)\}_{(q,p)} + \frac{\partial p'_i}{\partial t} \\ &= -\frac{\partial H'(q',p',t)}{\partial q'_i} \qquad = \{p'_i, H'(q',p',t)\}_{(q',p')}\,.\end{aligned}$$

The Poission bracket relations are preserved by canonical transformations:

$$\{q_i, p_j\}_{(q,p)} = \{q'_i, p'_j\}_{(q',p')} = \delta_{ij}\,.$$

Clearly, this cannot be done for any transformation. The restriction that we can find a new Hamiltonian $H'(q',p',t)$ such that the new coordinates satisfy Hamilton's equations is a severe one. We can prove that for transformations which can be obtained from a "generating function," it is always possible to find a new Hamiltonian. In order to show this, we start by constructing the Lagrangians in the two systems of coordinates:

$$\begin{aligned}L(q,\dot{q},t) &= p_i\dot{q}_i - H(q,p,t)\,, \\ L'(q',\dot{q}',t) &= p'_i\dot{q}'_i - H'(q',p',t)\,.\end{aligned} \tag{2.57}$$

Now $L(q,\dot{q},t)$ and $L'(q',\dot{q}',t)$ must satisfy Lagrange's equations in both systems, since Hamilton's equations are satisfied. Therefore they can differ by, at most, a total derivative of the coordinates q, q', and t:

$$L(q,\dot{q},t) = L'(q',\dot{q}',t) + \frac{dW(q,q',t)}{dt}\,. \tag{2.58}$$

So using (2.57) and (2.58), we find:

$$p_i = +\frac{\partial W(q,q',t)}{\partial q_i} = p_i(q,q',t)\,, \tag{2.59}$$

$$p'_i = -\frac{\partial W(q,q',t)}{\partial q'_i} = P_i(q,q',t)\,, \tag{2.60}$$

$$H'(q',p',t) = H(q,p,t) + \frac{\partial W(q,q',t)}{\partial t}\,. \tag{2.61}$$

Inverting (2.59) and (2.60) gives the canonical transformation (2.56), with the new Hamiltonian (2.61).

Now let the corresponding unitary transformation in quantum mechanics be $U(Q,P,t)$. Then the canonical transformation (2.56) is, in quantum mechanics, given by the unitary transformation,

$$\begin{aligned}Q'_i &= U^\dagger(Q,P,t)\,Q_i\,U(Q,P,t)\,, \\ P'_i &= U^\dagger(Q,P,t)\,P_i\,U(Q,P,t)\,.\end{aligned}$$

Everything here is in the Heisenberg picture. It is easy to show that

$$U(Q',P',t) = U(Q,P,t)\,.$$

So the eigenvectors of $Q(t)$ and $P(t)$ are transformed according to:

$$\begin{aligned}|\,q',t\,\rangle &= U^\dagger(Q,P,t)\,|\,q,t\,\rangle\,, \\ |\,p',t\,\rangle &= U^\dagger(Q,P,t)\,|\,p,t\,\rangle\,,\end{aligned}$$

So what we seek are the matrix elements,

$$U(q,q',t) = \langle\, q,t\,|\,q',t\,\rangle = \langle\, q,t\,|\,U^\dagger(Q,P,t)\,|\,q,t\,\rangle\,.$$

For the transformation generated by $W(q,q',t)$, we try to simply replace the classical function by a function of operators. This may result in ordering problems, which will need to be resolved in each case. Then from (2.59) and (2.60),

$$P_i = +\frac{\partial W(Q,Q',t)}{\partial Q_i}\,, \qquad P'_i = -\frac{\partial W(Q,Q',t)}{\partial Q'_i}\,,$$

$$H'(Q',P',t) - H(Q,P,t) = \frac{\partial W(Q,Q',t)}{\partial t}\,.$$

Therefore taking matrix elements of these three expressions between $\langle\, q,t\,|$ and $|\,q',t\,\rangle$, and using Schrödinger's equation of motion for $\langle\, q,t\,|$ and $|\,q',t\,\rangle$, gives:

$$\begin{aligned}
\langle\, q,t\,|\frac{\partial W(Q,Q',t)}{\partial Q_i}|\,q',t\,\rangle &= +\langle\, q,t\,|P_i|\,q',t\,\rangle = +\frac{\hbar}{i}\frac{\partial}{\partial q_i}\langle\, q,t\,|\,q',t\,\rangle\,,\\
\langle\, q,t\,|\frac{\partial W(Q,Q',t)}{\partial Q'_i}|\,q',t\,\rangle &= -\langle\, q,t\,|P'_i|\,q',t\,\rangle = +\frac{\hbar}{i}\frac{\partial}{\partial q'_i}\langle\, q,t\,|\,q',t\,\rangle\,,\\
\langle\, q,t\,|\frac{\partial W(Q,Q',t)}{\partial t}|\,q',t\,\rangle &= \langle\, q,t\,|\{\,H'(Q',P',t) - H(Q,P,t)\,\}|\,q',t\,\rangle\,,\\
&= \frac{\hbar}{i}\frac{\partial}{\partial t}\langle\, q,t\,|\,q',t\,\rangle\,.
\end{aligned}$$

Multiplying the first expression by δq_i, the second by $\delta q'_i$, the third by δt, and adding all three gives Schwinger's equation [**?**] for the transformation bracket:

$$\delta\langle\, q,t\,|\,q',t\,\rangle = \frac{i}{\hbar}\langle\, q,t\,|\,\delta W(Q(t),Q'(t),t)\,|\,q',t\,\rangle\,, \tag{2.62}$$

where the δ variation means:

$$\delta = \delta q_i\frac{\partial}{\partial q_i} + \delta q'_i\frac{\partial}{\partial q'_i} + \delta t\frac{\partial}{\partial t}\,.$$

A useful application of Schwinger's formula is when the classical transformation $W(q,q',t)$ is chosen such that the transformed Hamiltonian is identically zero. In this case, since $H'(q',p,t) = 0$, we find:

$$\dot q' = 0\,, \qquad \dot p' = 0\,.$$

Thus $q'(t) = q'$ and $p'(t) = p'$ are constants of the motion. The classical generator of this transformation is given by the solution of the Hamiltonian-Jacobi equation,

$$H(q,\frac{\partial W(q,q',t)}{\partial q},t) + \frac{\partial W(q,q',t)}{\partial t} = 0\,. \tag{2.63}$$

However, since q' is a constant of the motion, a formal solution of the Hamiltonian-Jacobi equation is given by the action, expressed in terms of the variables $q(t)$, q', and t. We can prove this by noting that in this case,

$$\frac{dW(q(t),q',t)}{dt} = \frac{\partial W}{\partial q_i}\dot q_i + \frac{\partial W}{\partial t} = p_i\dot q_i - H(q,p,t) = L(q(t),q',t)\,,$$

so

$$W(q(t),q',t) = \int_0^t L(q(t),q',t)\,\mathrm{d}t\,, \tag{2.64}$$

where the integration of (2.64) is along the classical path. So Schwinger's formula for this case becomes:

$$\delta\langle\, q,t\,|\,q'\,\rangle = \frac{i}{\hbar}\langle\, q,t\,|\,\delta\int_0^t L(Q(t),Q',t)\,\mathrm{d}t\,|\,q'\,\rangle\,. \tag{2.65}$$

This variational principle is the starting point for Schwinger's development of quantum mechanics. It relates the solution of the Hamilton-Jacobi equation to the quantum mechanical transfer matrix, $\langle\, q,t\,|\,q'\,\rangle$. Note, however, that the Lagrangian here is to be written as a function of Q' and the solution $Q(t)$, and not $Q(t)$ and $\dot{Q}(t)$. That is, to calculate the integral in (2.65), one needs to know the solution to the dynamics. For infinitesimal transformations such that $q = q' + \Delta q$, we can use this formula to find the infinitesimal transformation, but then one needs to sum the result over all paths in coordinate space. We do this in Chapter 3, where we discuss Feynman's path integral approach to quantum mechanics. Here, we illustrate the use of Schwinger's formula with several examples.

Example 21 (exchange of position and momentum)**.** As an example, we consider the time-independent classical canonical transformation generated by $W(q,q') = qq'$. Then

$$p = \frac{\partial W(q,q')}{\partial q} = q'\,,\qquad p' = -\frac{\partial W(q,q')}{\partial q} = -q\,. \tag{2.66}$$

Therefore this transformation sets $q' = p$ and $p' = -q$, that is, it interchanges q and $-p$. Using (2.62) we find:

$$\begin{aligned}\langle\, q\,|\delta W(Q,Q')|\,q'\,\rangle &= \langle\, q\,|\{\delta q\, Q' + \delta q'\, Q\}|\,q'\,\rangle\\ &= (\delta q\, q' + \delta q'\, q)\langle\, q\,|\,q'\,\rangle\\ &= \delta\langle\, q\,|\,q'\,\rangle\,.\end{aligned}$$

So the solution of this equation for $\langle\, q\,|\,q'\,\rangle$ is:

$$\langle\, q\,|\,q'\,\rangle = \langle\, q\,|\,p\,\rangle = \mathcal{N}e^{iqq'/\hbar} = \mathcal{N}e^{ipq/\hbar}\,.$$

The normalization is fixed by the requirement that

$$\langle\, q\,|\,q'\,\rangle = \int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi\hbar}\,\langle\, q\,|\,p\,\rangle\langle\, p\,|\,q'\,\rangle = |\mathcal{N}|^2\delta(q-q') \equiv \delta(q-q')\,.$$

Therefore $\mathcal{N} = 1$, and we find:

$$\langle\, q\,|\,p\,\rangle = e^{ipq/\hbar}\,,$$

in agreement with our previous result.

Example 22 (the free particle)**.** As an example of the use of the Hamilton-Jacobi solutions, Eq. (2.65), we consider first the free particle. Here we have

$$\begin{aligned}Q(t) &= Q' + \frac{P'}{m}t\,,\\ \dot{Q}(t) &= \frac{P}{m} = \frac{Q(t)-Q'}{t}\,,\\ [Q',Q(t)] &= \frac{i\hbar t}{m}\,.\end{aligned}$$

So we find

$$W(Q(t),Q',t) = \int_0^t L(Q(t'),Q',t')\,\mathrm{d}t' = \frac{1}{2}m\,\dot{Q}^2(t)\,t = \frac{m}{2t}(Q(t)-Q')^2\,.$$

Now we have:

$$\begin{aligned}
\frac{\partial W(Q(t),Q',t)}{\partial Q(t)} &= \frac{m}{t}\,(\,Q(t)-Q'\,)\,,\\
\frac{\partial W(Q(t),Q',t)}{\partial Q'} &= -\frac{m}{t}\,(\,Q(t)-Q'\,)\,,\\
\frac{\partial W(Q(t),Q',t)}{\partial t} &= -\frac{m}{2t^2}\,(\,Q(t)-Q'\,)^2 = -\frac{m}{2t^2}\,(\,Q^2(t)-Q(t)\,Q'-Q'\,Q(t)+Q'^2\,)\\
&= -\frac{m}{2t^2}\,(\,Q^2(t)-2\,Q(t)\,Q'+Q'^2-[\,Q',Q(t)\,]\,)\\
&= -\frac{m}{2t^2}\,(\,Q^2(t)-2\,Q(t)\,Q'+Q'^2\,)-\frac{\hbar}{i}\,\frac{1}{2t}\,.
\end{aligned}$$

Note in the last line, that we must find the partial derivative of $W(Q(t),Q',t)$ with respect to t, holding Q' and $Q(t)$ constant. Now that we have correctly ordered this expression, we can find the matrix elements needed to apply Eq. (2.65). We find:

$$\begin{aligned}
\frac{\partial\langle\,q,t\,|\,q'\,\rangle}{\partial q} &= +\frac{i}{\hbar}\frac{m}{2t}(q(t)-q')\,\langle\,q,t\,|\,q'\,\rangle\,,\\
\frac{\partial\langle\,q,t\,|\,q'\,\rangle}{\partial q'} &= -\frac{i}{\hbar}\frac{m}{2t}(q(t)-q')\,\langle\,q,t\,|\,q'\,\rangle\,,\\
\frac{\partial\langle\,q,t\,|\,q'\,\rangle}{\partial t} &= -\left\{\frac{i}{\hbar}\,\frac{m}{2t^2}(q-q')^2+\frac{1}{2t}\right\}\,\langle\,q,t\,|\,q'\,\rangle\,,
\end{aligned}$$

which has the solution,

$$\langle\,q,t\,|\,q'\,\rangle = \frac{\mathcal{N}}{\sqrt{t}}\exp\left\{\frac{i}{\hbar}\,\frac{m}{2t}(q-q')^2\right\}\,. \tag{2.67}$$

The normalization is fixed by the requirement that,

$$\lim_{t\to 0}\langle\,q,t\,|\,q'\,\rangle = \langle\,q\,|\,q'\,\rangle = \delta(q-q')\,.$$

A representation for the delta function is:

$$\lim_{\lambda\to 0^+}\frac{1}{\sqrt{\pi\lambda}}e^{-x^2/\lambda} = \delta(x)\,.$$

This gives $\mathcal{N}=\sqrt{m/2\pi i\hbar}$, so that

$$\langle\,q,t\,|\,q'\,\rangle = \sqrt{\frac{m}{2\pi i\hbar\,t}}\,\exp\left\{\frac{i}{\hbar}\,\frac{m}{2t}(q-q')^2\right\}\,. \tag{2.68}$$

Remark 11. We can also find the free particle transformation function directly, using the time-development operator. We find:

$$\begin{aligned}
\langle\,q,t\,|\,q',t'\,\rangle &= \langle\,q\,|\,U^\dagger(t-t')\,|\,q'\,\rangle = \langle\,q\,|\exp\left\{-\frac{i}{\hbar}\,\frac{P^2}{2m}\,(t-t')\right\}|\,q'\,\rangle\\
&= \int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi\hbar}\,\langle\,q\,|\,p\,\rangle\,\exp\left\{-\frac{i}{\hbar}\,\frac{p^2}{2m}\,(t-t')\right\}\langle\,p\,|\,q'\,\rangle\,,\\
&= \int_{-\infty}^{\infty}\frac{dp}{2\pi\hbar}\,\exp\frac{i}{\hbar}\left\{p\,(q-q')-\frac{p^2}{2m}\,(t-t')\right\}\\
&= \sqrt{\frac{m}{2\pi i\hbar\,(t-t')}}\,\exp\left\{\frac{i}{\hbar}\,\frac{m}{2}\,\frac{(q-q')^2}{(t-t')}\right\}\,\Theta(t-t')\,,
\end{aligned} \tag{2.69}$$

where we have done the last integral by completing the square, and assuming that $t>t'$ to converge the integral. What we have found here is the *retarded* propagator for a free particle.

2.4 Schwinger's transformation theory

Schwinger developed a quantum mechanics based on solving the variational equation (2.65) for the transition matrix element $\langle\, q,t\,|\,q',t'\,\rangle$, which we write here as:

$$\delta\langle\, q,t\,|\,q',t'\,\rangle = \frac{i}{\hbar}\langle\, q,t\,|\,\delta\int_{t'}^{t} L(Q(t),Q',t)\,\mathrm{d}t\,|\,q',t'\,\rangle\,. \tag{2.70}$$

Let us first note that the variational principle can be applied to *any* set of complete states at time t, for example, eigenvectors of the energy $|\,n,t\,\rangle$, since:

$$|\,n,t\,\rangle = \int \mathrm{d}q\,|\,q,t\,\rangle\,\langle\, q,t\,|\,n,t\,\rangle\,, \tag{2.71}$$

Eq. (2.70) can be written as:

$$\delta\langle\, n,t\,|\,n',t'\,\rangle = \frac{i}{\hbar}\langle\, n,t\,|\,\delta\int_{t'}^{t} L(Q(t),Q',t)\,\mathrm{d}t\,|\,n',t'\,\rangle\,. \tag{2.72}$$

Secondly, the variation of the action can include source terms in the action as well as coordinate terms. In order to illustrate this, let us study a one-dimentional harmonic oscillator with a driving force, where the Lagrangian is given by:

$$L = \frac{1}{2}\,m\,(\dot{Q}^2 - \omega^2\,Q^2) + F(t)\,Q\,, \tag{2.73}$$

where $F(t)$ is an external driving force. Then the variation of the action with respect to this driving force is given by:

$$\delta\langle\, n,t\,|\,n',t'\,\rangle = \frac{i}{\hbar}\langle\, n,t\,|\int_{t'}^{t} Q(t)\delta F(t)\,\mathrm{d}t\,|\,n',t'\,\rangle\,. \tag{2.74}$$

References

Chapter 3

Path integrals

Path integrals were invented by Feynman, as an alternative formulation of quantum theory. It appears that Feynman was trying to make sense of a remark in Dirac's book on quantum mechanics that an exponent of the classical Lagrangian was somehow equivalent to the transformation bracket $\langle\, q,t \,|\, q',t' \,\rangle$ in the Heisenberg representation. We have seen in Section 2.4 from Schwinger's action principle that the Lagrangian, rather than the Hamiltonian, is the correct weighting factor for variations of the Heisenberg transformation bracket. Sometime later, it was recognized that path integrals provided generating functionals for the Green functions needed for computation of the dynamics in the Heisenberg representation. Path integrals are seldom calculated directly and very few of them are known; however, extensive use of them is made in quantum field theory to prove various theorems. Subsequently path integrals were developed for quantities obeying Fermi statistics, as well as Bose statistics, using Grassmann anti-commuting variables. Feynman's original paper [1] was published in the Reviews of Modern Physics in 1948. He subsequently developed more material and published a book [2] in 1965. Other useful references are Schulman [3], and a more technical book by Rivers [4]. Numerous additional references can be found in these books.

We first consider quantum mechanics in one dimension. Recall that if $\psi(q,t)$ is the Schrödinger wave function at point (q,t) for the system, then we can write

$$\psi(q,t) = \langle\, q,t \,|\, \psi \,\rangle = \int \mathrm{d}q' \,\langle\, q,t \,|\, q',t' \,\rangle\, \langle\, q',t' \,|\, \psi \,\rangle = \int \mathrm{d}q' \,\langle\, q,t \,|\, q',t' \,\rangle\, \psi(q',t')\,, \qquad \text{for } t > t'. \tag{3.1}$$

where $\langle\, q,t \,|\, q',t' \,\rangle$ is in the Heisenberg representation. We found this propagator for a free particle in Section 22, but it is very difficult to find for other Hamiltonian systems. We will find it useful to have a general expression for this propagator when we study Green functions.

In this chapter, we will find a general expression for the Heisenberg bracket $\langle\, q,t \,|\, q',t' \,\rangle$ by splitting up the space-time path in many small intervals. For simplicity, we develop path integrals here for systems with one degree of freedom, the results in this chapter are readily generalized to systems with n degrees of freedom.

3.1 Space-time paths

Consider first a particle in one-dimension described by the Hamiltonian $H(q,p)$. We define an arbitrary path $q(t)$ in space-time from t' to t such that $q(t) = q$ and $q(t') = q'$, and set up an equal spaced time grid given by

$$t_i = i\Delta t\,, \qquad q_i = q(t_i)\,, \qquad i = 1,2,\ldots,n, \tag{3.2}$$

Then we write:

$$\langle\, q,t \,|\, q',t' \,\rangle = \int \mathrm{d}q_n \cdots \int \mathrm{d}q_2 \int \mathrm{d}q_1 \,\langle\, q,t \,|\, q_n,t_n \,\rangle \cdots \langle\, q_2,t_2 \,|\, q_1,t_1 \,\rangle \langle\, q_1,t_1 \,|\, q',t' \,\rangle\,, \tag{3.3}$$

where $t > t_n > t_{n-1} > \cdots > t_2 > t_1 > t'$. Now the bracket of the Heisenberg basis states is given by:

$$\langle\, q_{i+1}, t_{i+1} \,|\, q_i, t_i \,\rangle = \langle\, q_{i+1} \,|\, U(t_{i+1})\, U^\dagger(t_i) \,|\, q_i \,\rangle = \langle\, q_{i+1} \,|\, U(t_{i+1} - t_i) \,|\, q_i \,\rangle \tag{3.4}$$

So to first order in Δt, we find:

$$\begin{aligned}
\langle\, q_{i+1}, t_{i+1} \,|\, q_i, t_i \,\rangle &= \langle\, q_{i+1} \,|\, \exp\Big\{ -\frac{i}{\hbar}\, H(Q,P)\Delta t \Big\} \,|\, q_i \,\rangle = \exp\Big\{ -\frac{i}{\hbar}\, H\Big(q_i,\, \frac{\hbar}{i}\, \frac{\partial}{\partial q_{i+1}} \Big)\, \Delta t \Big\}\, \langle\, q_{i+1} \,|\, q_i \,\rangle \\
&= \exp\Big\{ -\frac{i}{\hbar}\, H\Big(q_i,\, \frac{\hbar}{i}\, \frac{\partial}{\partial q_{i+1}} \Big)\, \Delta t \Big\} \int \frac{\mathrm{d}p_i}{2\pi\hbar}\, \langle\, q_{i+1} \,|\, p_i \,\rangle \langle\, p_i \,|\, q_i \,\rangle \\
&= \exp\Big\{ -\frac{i}{\hbar}\, H\Big(q_i,\, \frac{\hbar}{i}\, \frac{\partial}{\partial q_{i+1}} \Big)\, \Delta t \Big\} \int \frac{\mathrm{d}p_i}{2\pi\hbar}\, \exp\Big\{ \frac{i}{\hbar}\, p_i (q_{i+1} - q_i) \Big\} \\
&= \int \frac{\mathrm{d}p_i}{2\pi\hbar}\, \exp\Big\{ \frac{i}{\hbar}\, \big[\, p_i \dot{q}_i - H(q_i, p_i) \,\big]\, \Delta t \Big\}\,.
\end{aligned} \tag{3.5}$$

Here we have used the the relation $\Delta q = q_{i+1} - q_i = \dot{q}_i \Delta t$, and the fact that

$$H\Big(q_i,\, \frac{\hbar}{i}\, \frac{\partial}{\partial q_{i+1}} \Big)\, \exp\Big\{ \frac{i}{\hbar}\, p_i (q_{i+1} - q_i) \Big\} = \exp\Big\{ \frac{i}{\hbar}\, p_i (q_{i+1} - q_i) \Big\} H(q_i, p_i)\,. \tag{3.6}$$

So we find (there is one less p integral):

$$\langle\, q, t \,|\, q', t' \,\rangle = \int \mathrm{d}q_n \cdots \int \frac{\mathrm{d}q_2\, \mathrm{d}p_2}{2\pi\hbar} \int \frac{\mathrm{d}q_1\, \mathrm{d}p_1}{2\pi\hbar}\, \exp\left\{ \frac{i}{\hbar} \sum_{i=1}^{n} \{\, p_i \dot{q}_i - H(q_i, p_i) \,\}\, \Delta t \right\}\,. \tag{3.7}$$

We define a path integral as the infinite limit of a sum over all possible paths in coordinate space.

$$\lim_{n\to\infty} \left\{ \int \mathrm{d}q_n \cdots \int \frac{\mathrm{d}q_2\, \mathrm{d}p_2}{2\pi\hbar} \int \frac{\mathrm{d}q_1\, \mathrm{d}p_1}{2\pi\hbar} \right\} = \int \frac{\mathrm{D}q\, \mathrm{D}p}{2\pi\hbar}\,. \tag{3.8}$$

In this limit, Eq. (3.7) becomes:

$$\begin{gathered}
\langle\, q, t \,|\, q', t' \,\rangle = \langle\, q \,|\, U(t - t') \,|\, q' \,\rangle = \langle\, q \,|\, e^{-iH(t-t')/\hbar} \,|\, q' \,\rangle = \int \frac{\mathrm{D}q\, \mathrm{D}p}{2\pi\hbar}\, e^{iS[q,p]/\hbar}\,, \\
\text{where:} \qquad S[q,p] = \int_t^{t'} \{\, p\, \dot{q} - H(q,p) \,\}\, \mathrm{d}t\,.
\end{gathered} \tag{3.9}$$

The path integral is over all paths $q(t)$ and $p(t)$ such that the end point are fixed: $q(t) = q$, $q(t') = q'$ and $p(t) = p$, $p(t') = p'$. The paths *do not* go backward in time. If $H(q,p)$ is of the form:

$$H(q,p) = \frac{p^2}{2m} + V(q)\,, \tag{3.10}$$

we can carry out the path integral over all p_i's and simplify the path integral. We find:

$$\langle\, q_{i+1}, t_{i+1} \,|\, q_i, t_i \,\rangle = \int_{-\infty}^{+\infty} \frac{\mathrm{d}p_i}{2\pi\hbar}\, e^{i[p_i \dot{q}_i - p_i^2/2m]\Delta t/\hbar} = \sqrt{\frac{m}{2\pi i\hbar\, \Delta t}}\, e^{i\frac{1}{2} m \dot{q}_i^2\, \Delta t/\hbar} = \sqrt{\frac{m}{2\pi i\hbar\, \Delta t}}\, e^{iL(q_i, \dot{q}_i)\Delta t/\hbar}\,. \tag{3.11}$$

where we have evaluated the integral using (3.78d). Putting this into the path integral, we find

$$\langle\, q, t \,|\, q', t' \,\rangle = \langle\, q \,|\, U(t - t') \,|\, q' \,\rangle = \mathcal{N} \int_{q(t')=q'}^{q(t)=q} \mathrm{D}q(t'')\, \exp\left\{ \frac{i}{\hbar} \int_{t'}^{t} L(q, \dot{q})\, \mathrm{d}t'' \right\} \qquad \text{for } t > t'. \tag{3.12}$$

Here the normalization factor is given by

$$\mathcal{N} = \lim_{n\to\infty}\Big[\,\frac{m}{2\pi i\hbar\,\Delta t}\,\Big]^{(n-1)/2}. \tag{3.13}$$

This limit is not well defined, so that the normalization of the path integral is usually determined in other ways. Often it is not needed, as we will see in the following sections of this chapter. Here $S[q]$ is the *classical* action integral for the path $q(t)$. The sum over paths in the integral is over *all* possible paths that go forward in time between fixed end points, not only the classical path. The paths must be continuous paths but need not be continuous in the derivative $\dot q(t)$, so they can be quite wild-looking.

For paths that go backward in time,

$$\langle\, q,t\,|\,q',t'\,\rangle = \langle\, q\,|\,U(t-t')\,|\,q'\,\rangle = \mathcal{N}\int_{q(t')=q'}^{q(t)=q}\mathrm{D}q(t'')\,\exp\Big\{-\frac{i}{\hbar}\int_t^{t'}L(q,\dot q)\,\mathrm{d}t''\Big\}\qquad \text{for } t<t'. \tag{3.14}$$

Eqs. (3.12) and (3.14) are the major results of Feynman, and can be thought of as a third way to quantize a classical system, equivalent in every way to Schrödinger's equation or the Heisenberg equations of motion. Planck's constant $\hbar$ appears here as a factor to make the exponent dimensionless.

3.2 Some path integrals

Evaluating path integrals presents a considerable challenge, and only two are known. Essentially the only way to do functional integrals is to break the integral into small intervals. We illustrate this with several examples in the following exercises

Exercise 2. Let us evaluate the Feynman path integral for a free particle in one dimension, where $L(q,\dot q) = m\dot q^2/2$. Let us split up the integral into small time steps, and put $\dot q_i\Delta t \approx q_{i+1}-q_i$. Then the path integral we want to evaluate is

$$\langle\, q,t\,|\,q',t'\,\rangle = \lim_{n\to\infty}\Big[\,\frac{m}{2\pi i\hbar\,\Delta t}\,\Big]^{(n-1)/2}\int\mathrm{d}q_n\cdots\int\mathrm{d}q_2\int\mathrm{d}q_1\,\exp\Big\{\frac{im}{2\hbar\Delta t}\sum_{i=0}^{n}(\,q_{i+1}-q_i)^2\Big\}, \tag{3.15}$$

where $q_0\equiv q'$ and $q_{n+1}\equiv q$. The integrals are actually easy to do using (3.78g), and we find the result (which is left as an exercise)

$$\langle\, q,t\,|\,q',t'\,\rangle = \frac{\mathcal{N}}{\sqrt{(t-t')}}\,\exp\Big\{\frac{i}{\hbar}\,\frac{m\,(q-q')^2}{2(t-t')}\Big\}, \tag{3.16}$$

in agreement with Eq. (2.69).

Exercise 3. A more difficult problem is to evaluate the path integral for a harmonic oscillator, where $L(q,\dot q) = m(\,\dot q^2+\omega^2q^2)/2$. In this case, we find

$$\langle\, q,t\,|\,q',t'\,\rangle = \frac{\mathcal{N}}{\sqrt{\sin[\,\omega(t-t')]}}\,\exp\Big\{\frac{i}{\hbar}\,\frac{m\omega}{2\sin[\,\omega(t-t')\,]}\,\big[\,(q^2+q'^2)\,\cos[\,\omega(t-t')\,]-2\,qq'\,\big]\Big\}. \tag{3.17}$$

Exercise 4. The path integral

$$\langle\,+\infty,+\infty\,|\,-\infty,-\infty\,\rangle = \mathcal{N}\int_{-\infty}^{+\infty}\mathrm{D}q\,\exp\Big\{\frac{i}{\hbar}\int_{-\infty}^{+\infty}j(t)\,q(t)\,\mathrm{d}t\Big\} = \mathcal{N}\,\delta[\,j\,], \tag{3.18}$$

is a δ-functional of $j(t)$, which is not well defined. Here $\mathcal{N}$ an (infinite) constant.

3.3 Matrix elements of coordinate operators

One of the most important uses of path integrals is to find matrix elements and Green functions of Heisenberg operators. For example, suppose we want to find the matrix element

$$\langle\, q,t\,|\,Q(t_i)\,|\,q',t'\,\rangle\,, \tag{3.19}$$

where $t > t_i > t'$. Then from (3.20), we have

$$\begin{aligned}\langle\, q,t\,|\,Q(t_i)\,|\,q',t'\,\rangle &= \int \mathrm{d}q_n \cdots \int \mathrm{d}q_2 \int \mathrm{d}q_1\, \langle\, q,t\,|\,q_n,t_n\,\rangle \cdots \langle\, q_{i+1},t_{i+1}\,|\,Q(t_i)\,|\,q_i,t_i\,\rangle \cdots \langle\, q_1,t_1\,|\,q',t'\,\rangle \\ &= \int \mathrm{d}q_n \cdots \int \mathrm{d}q_2 \int \mathrm{d}q_1\, q_i\, \langle\, q,t\,|\,q_n,t_n\,\rangle \cdots \langle\, q_2,t_2\,|\,q_1,t_1\,\rangle \langle\, q_1,t_1\,|\,q',t'\,\rangle\,.\end{aligned} \tag{3.20}$$

So in the limit, $n \to \infty$,

$$\langle\, q,t\,|\,Q(t_1)\,|\,q',t'\,\rangle = \mathcal{N} \int_{q(t')=q'}^{q(t)=q} \mathrm{D}q\, q(t_1)\, e^{\,i\int_{t'}^{t} L(q,\dot{q})\,\mathrm{d}t/\hbar}\,. \tag{3.21}$$

In a similar way, the expectation value of two Heisenberg coordinate operators at two different times is given by

$$\langle\, q,t\,|\,Q(t_1)\,Q(t_2)\,|\,q',t'\,\rangle = \mathcal{N} \int_{q(t')=q'}^{q(t)=q} \mathrm{D}q\, q(t_1)\, q(t_2)\, e^{\,i\int_{t'}^{t} L(q,\dot{q})\,\mathrm{d}t/\hbar}\,, \tag{3.22}$$

as long as $t > t_1 > t'$ and $t > t_2 > t'$. Now the time-ordered product is defined by

$$\mathcal{T}\{\,Q(t_1)\,Q(t_2)\,\} = Q(t_1)\,Q(t_2)\,\Theta(t_1-t_2) + Q(t_2)\,Q(t_1)\,\Theta(t_2-t_1)\,, \tag{3.23}$$

so from (3.22), we find

$$\langle\, q,t\,|\,\mathcal{T}\{\,Q(t_1)\,Q(t_2)\,\}\,|\,q',t'\,\rangle = \mathcal{N} \int_{q(t')=q'}^{q(t)=q} \mathrm{D}q\, q(t_1)\, q(t_2)\, e^{\,i\int_{t'}^{t} L(q,\dot{q})\,\mathrm{d}t/\hbar}\,, \tag{3.24}$$

since matrix elements of the operators for both time-ordering cases are given by the same path integral. The step functions then factor out of the integral and sum to one. Generalizing this equation to the case of the time-ordered product of any number of Heisenberg operators, we have:

$$\langle\, q,t\,|\,\mathcal{T}\{\,Q(t_1)\,Q(t_2)\cdots Q(t_n)\,\}\,|\,q',t'\,\rangle = \mathcal{N} \int_{q(t')=q'}^{q(t)=q} \mathrm{D}q\, q(t_1)\, q(t_2)\cdots q(t_n)\, e^{\,i\int_{t'}^{t} L(q,\dot{q})\,\mathrm{d}t/\hbar}\,, \tag{3.25}$$

for $t > t'$. Similarly for matrix elements of anti-time-ordered operators, we find

$$\langle\, q,t\,|\,\mathcal{T}^*\{\,Q(t_1)\,Q(t_2)\cdots Q(t_n)\,\}\,|\,q',t'\,\rangle = \mathcal{N} \int_{q(t')=q'}^{q(t)=q} \mathrm{D}q\, q(t_1)\, q(t_2)\cdots q(t_n)\, e^{-i\int_{t}^{t'} L(q,\dot{q})\,\mathrm{d}t/\hbar}\,, \tag{3.26}$$

for $t < t'$.

3.4 Generating functionals

Transitions amplitudes of time-ordered products can be obtained from a path integral generating functional. We introduce a classical external driving function $j_+(t)$ and define a generating functional $Z^{(+)}(q,t;q',t)[\,j_+\,]$

by a sum of functional derivatives

$$\begin{aligned} Z_{(+)}(q,t;q',t')[\,j_+\,] &= \sum_{n=0}^{\infty} \frac{1}{n!}\Big(\frac{i}{\hbar}\Big)^n \int_{t'}^{t} \mathrm{d}t_1 \int_{t'}^{t} \mathrm{d}t_2 \cdots \int_{t'}^{t} \mathrm{d}t_n \\ &\qquad \times \langle\, q,t \,|\, \mathcal{T}\{\, Q(t_1)\, Q(t_2) \cdots Q(t_n)\,\} \,|\, q',t' \,\rangle\, j_+(t_1)\, j_+(t_2) \cdots j_+(t_n) \\ &= \mathcal{N} \int_{q(t')=q'}^{q(t)=q} \mathrm{D}q\; e^{\,i\int_{t'}^{t} L(q,\dot q)\,\mathrm{d}t/\hbar} \sum_{n=0}^{\infty} \frac{1}{n!} \Big(\frac{i}{\hbar}\Big)^n \left[\, \int_{t'}^{t} \mathrm{d}t\, q(t)\, j_+(t) \right]^n \\ &= \mathcal{N} \int_{q(t')=q'}^{q(t)=q} \mathrm{D}q\; \exp\left\{ \frac{i}{\hbar} \int_{t'}^{t} [\, L(q,\dot q) + q(t)\, j_+(t)\,]\, \mathrm{d}t \right\}. \end{aligned} \tag{3.27}$$

But expanding $Z^{(+)}(q,t;q',t')[\,j_+\,]$ in a power series in j_+, we have

$$\begin{aligned} &Z_{(+)}(q,t;q',t')[\,j_+\,] \\ &\qquad = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{t'}^{t} \mathrm{d}t_1 \int_{t'}^{t} \mathrm{d}t_2 \cdots \int_{t'}^{t} \mathrm{d}t_n \left[\frac{\delta^n Z_{(+)}(q,t;q',t)[\,j\,]}{\delta j_+(t_1)\delta j_+(t_2)\cdots\delta j_+(t_n)} \right]_{j=0} j_+(t_1)\, j_+(t_2) \cdots j_+(t_n)\,. \end{aligned} \tag{3.28}$$

Comparing Eq. (3.27) with Eq. (3.28), we find

$$\langle\, q,t \,|\, \mathcal{T}\{\, Q(t_1)\, Q(t_2) \cdots Q(t_n)\,\} \,|\, q',t' \,\rangle = \Big(\frac{\hbar}{i}\Big)^n \left. \frac{\delta^n\, Z_{(+)}(q,t;q',t')[\,j\,]}{\delta j_+(t_1)\delta j_+(t_2)\cdots\delta j_+(t_n)} \right|_{j=0}, \tag{3.29}$$

for $t > t'$. So if we compute the generating functional $Z^{(+)}(q,t;q',t')[\,j_+\,]$, we can find expectation values for all the time-ordered products by functional differentiation.

In a similar way, the generating functional for anti-time-ordered products is given by

$$Z_{(-)}(q,t;q',t')[\,j_-\,] = \mathcal{N} \int_{q(t')=q'}^{q(t)=q} \mathrm{D}q\; \exp\left\{ -\frac{i}{\hbar} \int_{t}^{t'} [\, L(q,\dot q) + q(t)\, j_+(t)\,]\, \mathrm{d}t \right\}, \tag{3.30}$$

from which we find matrix elements of all the anti-time-ordered products

$$\langle\, q,t \,|\, \mathcal{T}^*\{\, Q(t_1)\, Q(t_2) \cdots Q(t_n)\,\} \,|\, q',t' \,\rangle = \Big(-\frac{\hbar}{i}\Big)^n \left. \frac{\delta^n\, Z_{(-)}(q,t;q',t')[\,j_-\,]}{\delta j_-(t_1)\delta j_-(t_2)\cdots\delta j_-(t_n)} \right|_{j_-=0}, \tag{3.31}$$

for $t < t'$.

3.5 Closed time path integrals

Time-ordered products are useful in problems when we know the state of the system in the distant future or past. However for initial value problems where we want to find the density matrix as a function of time, we will need two propagators and two path integrals, one which propagates forward in time from a point q'' at $t = 0$ to a point q at time t and then one which propagates backward in time from a point q' at time t to a point q at time $t = 0$, as explained in the introduction to this chapter.

The use of closed-time-path Green functions was first done by Schwinger [5], later by Keldysh [6], and further developed by Bakshi and Mahanthappa [7, 8].

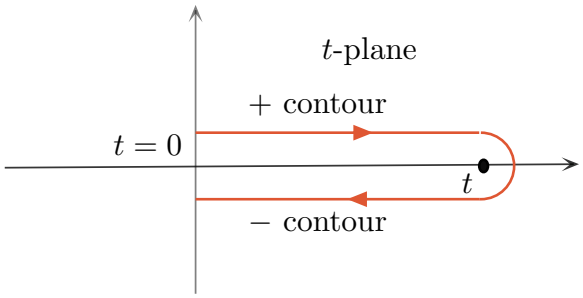

Figure 3.1: The closed time path contour.

From Eq. (**??**), the density matrix $\rho(q,q',t)$ at time t is given by

$$\begin{aligned}\rho(q,q',t) &= \iint_{-\infty}^{+\infty} \mathrm{d}q''\,\mathrm{d}q'''\,\rho_0(q'',q''')\,\langle\, q,t\,|\,q'',0\,\rangle\,\langle\, q''',0\,|\,q',t\,\rangle \\ &= \mathcal{N}\iint_{-\infty}^{+\infty} \mathrm{d}q''\,\mathrm{d}q'''\,\rho_0(q'',q''')\int_{q(0)=q''}^{q(t)=q}\mathrm{D}q_+\int_{q(t)=q'}^{q(0)=q'''}\mathrm{D}q_- \\ &\quad\times\exp\left\{\frac{i}{\hbar}\int_0^t L(q_+,\dot q_+)\,\mathrm{d}t' - \int_t^0 L(q_-,\dot q_-)\,\mathrm{d}t'\right\}\end{aligned} \tag{3.32}$$

Normalization of the state vector requires

$$\begin{aligned}\langle\,\Psi(t)\,|\,\Psi(t)\,\rangle &= \int_{-\infty}^{+\infty}\mathrm{d}q\,\langle\,\Psi(t)\,|\,q\,\rangle\,\langle\,q\,|\,\Psi(t)\,\rangle = \int_{-\infty}^{+\infty}\mathrm{d}q\,\langle\,q\,|\,\rho(t)\,|\,q\,\rangle = \int_{-\infty}^{+\infty}\mathrm{d}q\,\rho(q,q,t) \\ &= \mathcal{N}\iiint_{-\infty}^{+\infty}\mathrm{d}q\,\mathrm{d}q''\,\mathrm{d}q'''\,\rho_0(q'',q''')\int_{q(0)=q''}^{q(t)=q}\mathrm{D}q_+\int_{q(t)=q}^{q(0)=q'''}\mathrm{D}q_- \\ &\quad\times\exp\left\{\frac{i}{\hbar}\int_0^t L(q_+,\dot q_+)\,\mathrm{d}t' - \int_t^0 L(q_-,\dot q_-)\,\mathrm{d}t'\right\} \\ &= \mathcal{N}\iint_{-\infty}^{+\infty}\mathrm{d}q''\,\mathrm{d}q'''\,\rho_0(q'',q''')\int_{q(0)=q''}^{q(0)=q'''}Dq\exp\left\{\frac{i}{\hbar}\int_{\mathcal{C}}L(q,\dot q)\,\mathrm{d}t'\right\}\equiv 1\,,\end{aligned} \tag{3.33}$$

which fixes the normalization factor,

$$1/\mathcal{N} = \iint_{-\infty}^{+\infty}\mathrm{d}q''\,\mathrm{d}q'''\,\rho_0(q'',q''')\int_{q(0)=q''}^{q(0)=q'''}Dq\exp\left\{\frac{i}{\hbar}\int_{\mathcal{C}}L(q,\dot q)\,\mathrm{d}t'\right\}. \tag{3.34}$$

Here the path integral goes from q'' at $t=0$ to all points q at t and then back to q''' at $t=0$. We have included the integral over $\mathrm{d}q$ in the path integral. The time integral for the action goes over a *closed time path contour* $\mathcal{C}$ in the complex t-plane shown in Fig. 3.1. The integral goes from $t=0$ a distance ϵ above the real t-axis up to an arbitrary time t and back a distance ϵ below the real t-axis to $t=0$. We then take the limit $\epsilon\to 0$ and define $q(t+i\epsilon)=q_+(t)$ and $q(t-i\epsilon)=q_-(t)$, so that both $q_\pm(t)$ are included in the integral. Note that the path below the real axis is in the *negative* real t-direction.

Now let us find the average value of the Heisenberg operator $Q(t)$ at time t for the initial state $|\,\Psi\,\rangle$. This

is given by

$$\begin{aligned}\langle\, Q(t)\,\rangle &= \langle\, \Psi\,|\,Q(t)\,|\,\Psi\,\rangle = \langle\, \Psi(t)\,|\,Q\,|\,\Psi(t)\,\rangle = \int_{-\infty}^{+\infty} \mathrm{d}q\, q\, \rho(q,q,t) \\ &= \mathcal{N}\iint_{-\infty}^{+\infty} \mathrm{d}q''\,\mathrm{d}q'''\,\rho_0(q'',q''') \int_{q(0)=q''}^{q(0)=q'''} Dq\, q(t)\, \exp\left\{\frac{i}{\hbar}\int_{\mathcal{C}} L(q,\dot{q})\,\mathrm{d}t'\right\}.\end{aligned} \tag{3.35}$$

Here $q(t)$ is evaluated at the point t of the closed time path, on either the upper or lower branch. From our analysis in Section 3.3, the path integral for the expectation value of two Heisenberg operators $\langle\, Q(t_1)\,Q(t_2)\,\rangle$ is given by

$$\langle\, Q(t_1)\,Q(t_2)\,\rangle = \mathcal{N}\iint_{-\infty}^{+\infty} \mathrm{d}q''\,\mathrm{d}q'''\,\rho_0(q'',q''') \int_{q(0)=q''}^{q(0)=q'''} Dq\, q(t_1)\, q(t_2)\, \exp\left\{\frac{i}{\hbar}\int_{\mathcal{C}} L(q,\dot{q})\,\mathrm{d}t'\right\}. \tag{3.36}$$

In this case, however, we can evaluate $q(t_1)$ and $q(t_2)$ in four different ways, depending on which branch of the closed-time-path contour they are on. Keeping in mind the *direction* of time as shown in the closed time path of Fig. 3.1, we define:

$$\Theta_{\mathcal{C}}(t_1,t_2) = \begin{cases} 1\,, & \text{for } t_1 \text{ later than } t_2, \\ 0\,, & \text{for } t_1 \text{ earlier than } t_2. \end{cases} \tag{3.37}$$

Recall that all times on the *lower* branch are *later* and run backward than those on the upper branch!

Exercise 5. Show that

$$\Theta_{\mathcal{C}}(t_1,t_2) = \begin{cases} \Theta(t_1-t_2)\,, & \text{for } t_1 \text{ on upper contour and } t_2 \text{ on upper contour}, \\ 1\,, & \text{for } t_1 \text{ on upper contour and } t_2 \text{ on lower contour}, \\ 0\,, & \text{for } t_1 \text{ on lower contour and } t_2 \text{ on upper contour}, \\ \Theta(t_2-t_1)\,, & \text{for } t_1 \text{ on lower contour and } t_2 \text{ on lower contour}. \end{cases} \tag{3.38}$$

So let us define the closed-time-path ordering by

$$\mathcal{T}_{\mathcal{C}}\{\, Q(t_1)\,Q(t_2)\,\} = Q(t_1)\,Q(t_2)\,\Theta_{\mathcal{C}}(t_1,t_2) + Q(t_2)\,Q(t_1)\,\Theta_{\mathcal{C}}(t_2,t_1)\,. \tag{3.39}$$

Then we have

$$\langle\, \mathcal{T}_{\mathcal{C}}\{\, Q(t_1)\,Q(t_2)\,\}\,\rangle = \mathcal{N}\iint_{-\infty}^{+\infty} \mathrm{d}q''\,\mathrm{d}q'''\,\rho_0(q'',q''') \int_{q(0)=q''}^{q(0)=q'''} Dq\, q(t_1)\, q(t_2)\, \exp\left\{\frac{i}{\hbar}\int_{\mathcal{C}} L(q,\dot{q})\,\mathrm{d}t'\right\}, \tag{3.40}$$

where $q(t_1)$ and $q(t_2)$ are evaluated on the closed time path contour. We can generalize this to expectation values of any number of closed-time-path time-ordered products

$$\begin{aligned}&\langle\, \mathcal{T}_{\mathcal{C}}\{\, Q(t_1)\,Q(t_2)\cdots Q(t_n)\,\}\,\rangle \\ &\qquad = \mathcal{N}\iint_{-\infty}^{+\infty} \mathrm{d}q''\,\mathrm{d}q'''\,\rho_0(q'',q''') \int_{q(0)=q''}^{q(0)=q'''} Dq\, q(t_1)\, q(t_2)\cdots q(t_n)\, \exp\left\{\frac{i}{\hbar}\int_{\mathcal{C}} L(q,\dot{q})\,\mathrm{d}t'\right\}.\end{aligned} \tag{3.41}$$

Now we are in a position to find a generating function for these closed-time-path Green functions. Following

our methods in Section 3.4, we define a generating functional $Z_{\mathcal{C}}[\,j\,]$ by

$$\begin{aligned} Z_{\mathcal{C}}[\,j\,] &= \sum_{n=0}^{\infty} \frac{1}{n!}\Big(\frac{i}{\hbar}\Big)^n \int_{\mathcal{C}} \mathrm{d}t_1 \int_{\mathcal{C}} \mathrm{d}t_2 \cdots \int_{\mathcal{C}} \mathrm{d}t_n \\ &\qquad \times \langle\, \mathcal{T}_{\mathcal{C}}\{\, Q(t_1)\, Q(t_2) \cdots Q(t_n)\,\}\,\rangle\, j(t_1)\, j(t_2) \cdots j(t_n) \\ &= \mathcal{N} \iint_{-\infty}^{+\infty} \mathrm{d}q''\, \mathrm{d}q'''\, \rho_0(q'', q''') \int_{q(0)=q''}^{q(0)=q'''} Dq\; e^{i \int_{\mathcal{C}} L(q,\dot{q})\, \mathrm{d}t'/\hbar} \sum_{n=0}^{\infty} \frac{1}{n!} \Big(\frac{i}{\hbar}\Big)^n \left[\int_{\mathcal{C}} q(t')\, j(t')\, \mathrm{d}t' \right]^n \\ &= \mathcal{N} \iint_{-\infty}^{+\infty} \mathrm{d}q''\, \mathrm{d}q'''\, \rho_0(q'', q''') \int_{q(0)=q''}^{q(0)=q'''} Dq\; \exp\left\{ \frac{i}{\hbar} \int_{\mathcal{C}} [\, L(q, \dot{q}) + q(t')\, j(t')\,]\, \mathrm{d}t' \right\}. \end{aligned} \tag{3.42}$$

Using the normalization factor given in Eq. (3.34), the generating functional is normalized so that $Z_{\mathcal{C}}[j = 0] = 1$. Then

$$\langle\, \mathcal{T}_{\mathcal{C}}\{\, Q(t_1)\, Q(t_2) \cdots Q(t_n)\,\}\,\rangle = \Big(\frac{\hbar}{i}\Big)^n \left. \frac{\delta^n\, Z_{\mathcal{C}}[\,j\,]}{\delta j(t_1) \delta j(t_2) \cdots \delta j(t_n)} \right|_{j=0} . \tag{3.43}$$

The average value of $Q(t)$ is given by

$$q(t) = \langle\, \mathcal{T}_{\mathcal{C}}\{\, Q(t)\,\}\,\rangle = \left(\frac{\hbar}{i}\right) \left. \frac{\delta Z_{\mathcal{C}}[\,j\,]}{\delta j(t)} \right|_{j=0} \tag{3.44}$$

The closed-time-path Green function is defined by

$$G(t, t') = i\, \langle\, \mathcal{T}_{\mathcal{C}}\{\, Q(t)\, Q(t')\,\}\,\rangle/\hbar = \left(\frac{\hbar}{i}\right) \left. \frac{\delta^2 Z_{\mathcal{C}}[\,j\,]}{\delta j(t)\, \delta j(t')} \right|_{j=0} . \tag{3.45}$$

There are *four* such Green functions, depending on where we evaluate t and t' on the closed-time-path contour. These are the Green functions we used in previous chapters to study problems in the Heisenberg representation.

Exercise 6. Writing the four closed-time-path contour Green functions in a matrix notation, show that

$$\begin{aligned} G^{ab}(t, t') &= \frac{i}{\hbar} \begin{pmatrix} \langle\, \mathcal{T}\{\, Q(t)\, Q(t')\,\}\,\rangle & \langle\, Q(t)\, Q(t')\,\rangle \\ \langle\, Q(t')\, Q(t)\,\rangle & \langle\, \mathcal{T}^*\{\, Q(t)\, Q(t')\,\}\,\rangle \end{pmatrix} \\ &= G_{>}(t, t')\, \Theta_c^{ab}(t, t') + G_{<}(t, t')\, \Theta_c^{ab}(t', t)\,, \end{aligned} \tag{3.46}$$

where

$$\begin{aligned} \langle\, \mathcal{T}\{\, Q(t)\, Q(t')\,\}\,\rangle &= \langle\, Q(t)\, Q(t')\,\rangle\, \Theta(t - t') + \langle\, Q(t')\, Q(t)\,\rangle\, \Theta(t' - t)\,, \\ \langle\, \mathcal{T}^*\{\, Q(t)\, Q(t')\,\}\,\rangle &= \langle\, Q(t')\, Q(t)\,\rangle\, \Theta(t - t') + \langle\, Q(t)\, Q(t')\,\rangle\, \Theta(t' - t)\,, \end{aligned} \tag{3.47}$$

and

$$\Theta_c^{ab}(t, t') = \begin{pmatrix} \Theta(t - t') & 0 \\ 1 & \Theta(t' - t) \end{pmatrix}, \qquad \Theta_c^{ab}(t', t) = \begin{pmatrix} \Theta(t' - t) & 1 \\ 0 & \Theta(t - t') \end{pmatrix}. \tag{3.48}$$

and where we have put

$$G_{>}(t, t') = i\langle\, Q(t)\, Q(t')\,\rangle/\hbar\,, \qquad G_{<}(t, t') = i\langle\, Q(t')\, Q(t)\,\rangle/\hbar\,. \tag{3.49}$$

3.6 Initial value conditions

We have yet to set initial conditions for the Green functions. These can be fixed by the initial density matrix, which we can always write as an exponent of a power series in q and q'. We put

$$\rho(q,q') = e^{i\Phi(q,q')/\hbar}\,, \qquad \text{where} \qquad \Phi(q,q') = J_0 + J_{1,0}\,q + J_{0,1}\,q' + J_{1,1}\,q\,q' + \cdots\,, \tag{3.50}$$

where $J_{i,j}$ are constants, fixed by the initial density matrix. Here q and q' are the end points of the path integral at $t = 0$. We can incorporate these initial density matrix expansion terms into the Lagrangian as multi-point currents with support only at $t = 0$. That is, we put

$$\Phi(q,q') = \int_{\mathcal{C}} \phi(t)\,\mathrm{d}t\,, \quad \text{where} \quad \phi(t) = \sum_{i,j} j_{i,j}(t)\,q^i(t)\,q'^j(t)\,, \quad \text{with} \quad j_{i,j}(t) = J_{i,j}\,\delta(t)\,. \tag{3.51}$$

where $j_{i,j}(t) = J_{i,j}\,\delta(t)$. So we can redefine the Lagrangian to include these extra current terms, $L'(t) = L(t) + \phi(t)$, and simplify the generating functional to give the expression

$$Z_{\mathcal{C}}[\,j\,] = \mathcal{N}\int \mathrm{D}q\; e^{iS[\,q,j\,]/\hbar}\,, \qquad \text{where} \qquad S[\,q,j\,] = \int_c [\,L'(q,\dot{q}) + q\,j\,]\,\mathrm{d}t'\,. \tag{3.52}$$

We will see in the next section that these extra currents and driving forces do not effect the equations of motion for $t \neq 0$, and only serve to provide initial conditions for the vertex functions.

3.7 Connected Green functions

A generator $W[\,j\,]$ for connected Green functions are defined by

$$Z[\,j\,] = e^{iW[\,j\,]/\hbar}\,. \tag{3.53}$$

Now let us define

$$q[j](t) = \left(\frac{\hbar}{i}\right)\,\frac{1}{Z[j]}\,\frac{\delta Z[\,j\,]}{\delta j(t)} = \frac{\delta W[\,j\,]}{\delta j(t)}\,, \tag{3.54}$$

which is a functional of j. We assume that we can invert this expression to find $j(t)$ as a function of $q(t)$. For the two-point functions, we find

$$G[\,j\,](t,t') = \left(\frac{\hbar}{i}\right)\,\frac{1}{Z[j]}\,\frac{\delta^2 Z[\,j\,]}{\delta j(t)\,\delta j(t')} = W[\,j\,](t,t') + \left(\,\frac{i}{\hbar}\right) q(t)\,q(t)\,, \tag{3.55}$$

where

$$W[\,j\,](t,t') = \frac{\delta^2 W[\,j\,]}{\delta j(t)\,\delta j(t')}\,, \tag{3.56}$$

which is also a functional of j. We will now find it useful to define a vertex function $\Gamma[\,q\,](t,t')$, which is a functional of $q(t)$ by a Legendre transformation,

$$\Gamma[\,q\,] = \int_{\mathcal{C}} \mathrm{d}t\,q(t)\,j(t) - W[\,j\,]\,. \tag{3.57}$$

Then

$$j(t) = \frac{\delta\Gamma[\,q\,]}{\delta q(t)}\,. \tag{3.58}$$

Vertex functions are defined by multiple derivatives of $\Gamma[\,q\,]$ with respect to $q(t)$. For example, the two-point vertex function $\Gamma_{ab}[\,q\,](t,t')$ is defined by

$$\Gamma[\,q\,](t,t') = \frac{\delta^2\Gamma[\,q\,]}{\delta q(t)\,\delta q(t')}\,. \tag{3.59}$$

The vertex function $\Gamma_{ab}[\,q\,](t,t')$ and the connected Green function $W^{ab}[\,j\,](t,t')$ are inverses of each other. We find

$$\int_{\mathcal{C}} \mathrm{d}t'\, \Gamma[\,q\,](t,t')\, \Gamma[\,q\,](t',t'') = \delta_{\mathcal{C}}(t,t'')\,, \tag{3.60}$$

where the closed-time-path delta function is defined as the derivative of the closed-time-path step function defined in Eq. (3.37).

3.8 Classical expansion

In order to find the path integral, we must sum over all paths from a point q' at time t' to a point q at time t; however one might suppose that for some problems, the most probable path would be the classical path, and that for such problems, a good approximation of the path integral would be the classical path plus small variations about the classical path. We look at such an approximate scheme in this section. We can think of this approximation as a limit as $\hbar \to 0$ so that we can use a method of steepest descent to evaluate the path integral, much like the WKB approximation for the Schrödinger equation.

We start by expanding the action given in Eq. (3.52) about a value $q_c(t)$,

$$S[\,q,j\,] = S[\,q_c,j\,] + \int_{\mathcal{C}} \mathrm{d}t\; \frac{\delta S[\,q,j\,]}{\delta q(t)}\bigg|_{q_c} (q(t)-q_c(t)) + \int_{\mathcal{C}}\mathrm{d}t \int_{\mathcal{C}}\mathrm{d}t'\; \frac{\delta^2 S[\,q,j\,]}{\delta q(t)\,\delta q(t')}\bigg|_{q_c} (q(t)-q_c(t))\,(q(t')-q_c(t')) + \cdots \tag{3.61}$$

Setting the first variation equal to zero

$$\frac{\delta S[\,q,j\,]}{\delta q(t)} = 0\,, \tag{3.62}$$

yields the classical Lagrange equations of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L(q,\dot q)}{\partial \dot q} - \frac{\partial L(q,\dot q)}{\partial q} = j\,, \tag{3.63}$$

which are to be evaluated at $q(t) = q_c[j](t)$. $q_c[j](t)$ is to be regarded as a functional of j. Then the first term $S[\,q_c,j\,]$ in (3.61) is just the classical action, which is a functional of j and comes out of the path integral. So we are left with the following expression for the generating functional

$$Z_{\mathcal{C}}[\,j\,] = \mathcal{N} e^{iS[\,q_c,j\,]/\hbar} \int \mathrm{D}q\; \exp\left\{ \frac{i}{\hbar}\int_{\mathcal{C}}\mathrm{d}t \int_{\mathcal{C}}\mathrm{d}t'\,(q(t)-q_c(t))\,\gamma(t,t')\,(q(t')-q_c(t')) + \cdots \right\}, \tag{3.64}$$

where

$$\gamma(t,t') = \frac{\delta^2 S[\,q,j\,]}{\delta q(t)\,\delta q(t')}\bigg|_{q_c}\,. \tag{3.65}$$

The quadratic path integral in Eq. (3.64) can easily be done. We first change variables by setting $q'(t) = q(t) - q_c(t)$. Then we break the integral into finite pieces. This gives

$$\mathcal{N}\int_{-\infty}^{+\infty}\mathrm{d}q_1' \int_{-\infty}^{+\infty}\mathrm{d}q_2' \cdots \int_{-\infty}^{+\infty}\mathrm{d}q_n'\; \exp\left\{ \frac{i}{2\hbar}\sum_{i,j=1}^{n} q_i'\,\gamma_{ij}\,q_j' \right\} \tag{3.66}$$

Next, we assume that we can bring γ_{ij} to diagonal form by a unitary transformation, $\gamma_{ij} = U^{\dagger}_{ik}\,\gamma'_k\,U_{kj}$ and

we define new variables $q_i'' = U_{ij} q_j'$. Then (3.66) becomes

$$\mathcal{N}\int_{-\infty}^{+\infty} \mathrm{d}q_1'' \int_{-\infty}^{+\infty} \mathrm{d}q_2'' \cdots \int_{-\infty}^{+\infty} \mathrm{d}q_n'' \, \exp\left\{ \frac{i}{2\hbar} \sum_k \gamma_k' \, q''2_k \right\}$$
$$= \frac{\mathcal{N}'}{\sqrt{\prod_k \gamma_k'}} = \frac{\mathcal{N}'}{\sqrt{\det[\,\gamma'\,]}} = \frac{\mathcal{N}'}{\sqrt{\det[\,U^\dagger \gamma U\,]}} = \frac{\mathcal{N}'}{\sqrt{\det[\,\gamma\,]}} = \mathcal{N}' \, \exp\left\{ -\frac{1}{2} \, \mathrm{Tr}[\,\ln[\,\gamma\,]\,] \right\}$$
$$\to \mathcal{N}' \, \exp\left\{ -\frac{1}{2} \int_{\mathcal{C}} \mathrm{d}t \, \ln[\,\gamma(t,t)\,] \right\} . \quad (3.67)$$

Adding this to the first term, we find that the expansion of the generating functional can be written as

$$Z_{\mathcal{C}}[\,j\,] = e^{iS_{\mathrm{eff}}[j]/\hbar} \, , \quad (3.68)$$

where the effective action is given by the expansion

$$S_{\mathrm{eff}}[j] = S_0 + S[\,q_c, j\,] + \frac{i\hbar}{2} \int_{\mathcal{C}} \mathrm{d}t \, \ln\{\,\gamma[j](t,t)\,\} + \cdots \, , \quad (3.69)$$

where S_0 is a constant and $q_c(t)$ is the solution of the classical equations of motion. Here $\gamma[j](t,t)$ is a functional of j. So the generating function is the classical action plus a trace-log term, which is proportional to $\hbar$ and is therefore a quantum effect. Comparing the definition of the generator of connected Green functions $W[\,j\,]$ in Eq. (3.56) with Eq. (3.68), we see that the effective action is just this generator

$$W[\,j\,] = S_{\mathrm{eff}}[j] \, . \quad (3.70)$$

This enables us to construct vertex function by the Lagendre transformation (3.57)

$$\Gamma[\,q\,] = \int_{\mathcal{C}} \mathrm{d}t \, q(t) \, j(t) - S_{\mathrm{eff}}[j] \, . \quad (3.71)$$

Example 23. Let us take a (not so simple) example, and work out the Green and vertex functions explicitly. The action $S[\,q, j\,]$ for an anharmonic oscillator is of the form

$$\begin{aligned} S[\,q, j\,] &= \int_{\mathcal{C}} \mathrm{d}t \left\{ \frac{m}{2} \, [\,\dot{q}^2(t) + \omega^2 \, q^2(t)\,] + \frac{\lambda}{4} \, q^4(t) + q(t) j(t) \right\} \\ &= \int_{\mathcal{C}} \mathrm{d}t \left\{ \frac{m}{2} \, q(t) \left[-\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega^2 \right] q(t) + \frac{\lambda}{4} \, q^4(t) + q(t) j(t) \right\} . \end{aligned} \quad (3.72)$$

Here we have integrated the kinetic energy term by parts and discarded the integrated factor at the end points of the closed-time-path integral. The first derivative of the action gives

$$\frac{\delta S[\,q, j\,]}{\delta q(t)} = m \left[-\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega^2 \right] q(t) + \lambda \, q^3(t) + j(t) \quad (3.73)$$

Setting this equal to zero gives a differential equation for $q_c(t)$ in terms of the current $j(t)$,

$$m \left[\frac{\mathrm{d}^2}{\mathrm{d}t^2} - \omega^2 \right] q_c(t) - \lambda \, q_c^3(t) = j(t) \, . \quad (3.74)$$

The second functional derivative with respect to $q(t)$ gives

$$\gamma(t,t') = \frac{\delta^2 S[\,q, j\,]}{\delta q(t) \, \delta q(t')} \bigg|_{q_c} = \left\{ m \left[-\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega^2 \right] + 3\lambda \, q_c^2(t) \right\} \delta_{\mathcal{C}}(t,t') \, , \quad (3.75)$$

which is a differential operator. It's inverse is a Green function. Differenting Eq. (3.74) with respect to $j(t')$ gives

$$\left\{ m\left[\frac{\mathrm{d}^2}{\mathrm{d}t^2} - \omega^2\right] - 3\lambda\, q_c^2(t)\right\} \frac{\delta q_c(t)}{\delta j(t')} = \delta_{\mathcal{C}}(t,t')\,. \tag{3.76}$$

So if we put $g(t,t') = \delta q_c(t)/\delta j(t')$, Eq. (3.76) states that

$$\int_{\mathcal{C}} \mathrm{d}t'\, \gamma(t,t')\, g(t',t'') = \delta_{\mathcal{C}}(t,t'')\,. \tag{3.77}$$

That is $g(t',t'')$ is the inverse of $\gamma(t,t')$.

3.9 Some useful integrals

Some useful integrals are

$$\int_{-\infty}^{+\infty} \mathrm{d}x\, e^{-ax^2} = \sqrt{\frac{a}{\pi}}\,, \tag{3.78a}$$

$$\int_{-\infty}^{+\infty} \mathrm{d}x\, x^2\, e^{-ax^2} = \sqrt{\frac{a}{\pi}}\left(\frac{1}{2a}\right), \tag{3.78b}$$

$$\int_{-\infty}^{+\infty} \mathrm{d}x\, x^4\, e^{-ax^2} = \sqrt{\frac{a}{\pi}}\left(\frac{3}{4a^2}\right), \tag{3.78c}$$

$$\int_{-\infty}^{+\infty} \mathrm{d}x\, e^{-ax^2+bx} = \sqrt{\frac{\pi}{a}}\, e^{b^2/4a}\,, \tag{3.78d}$$

$$\int_{-\infty}^{+\infty} \mathrm{d}x\, x\, e^{-ax^2+bx} = \sqrt{\frac{\pi}{a}}\left(\frac{b}{2a}\right) e^{b^2/4a}\,, \tag{3.78e}$$

$$\int_{-\infty}^{+\infty} \mathrm{d}x\, x^2\, e^{-ax^2+bx} = \sqrt{\frac{\pi}{a}}\left(\frac{1}{2a} + \frac{b^2}{4a^2}\right) e^{b^2/4a}\,, \tag{3.78f}$$

$$\int_{-\infty}^{+\infty} \mathrm{d}y \sqrt{\frac{a}{\pi}}\, e^{-a\,(x-y)^2} \sqrt{\frac{b}{\pi}}\, e^{-a\,(y-z)^2} = \sqrt{\frac{ab}{\pi(a+b)}}\, e^{-ab\,(x-z)^2/(a+b)}\,, \tag{3.78g}$$

References

[1] R. P. Feynman, "Space-time approach to non-relativistic quantum mechanics," Rev. Mod. Phys. **20**, 367 (1948).

[2] R. P. Feynman and A. R. Hibbs, *Quantum mechanics and path integrals* (McGraw-Hill, New York, 1965).

[3] L. S. Schulman, *Techniques and applications of path integration* (John Wiley & Sons, New York, 1981).

[4] R. J. Rivers, *Path integral methods in quantum field theory* (Cambridge University Press, 1990).

[5] J. Schwinger, "Brownian motion of a quantum oscillator," J. Math. Phys. **2**, 407 (1961).

[6] L. V. Keldysh, "Diagram technique for nonequilibrium processes," Zh. Eksp. Teor. Fiz. **47**, 1515 (1964). (Sov. Phys. JETP 20:1018,1965).

[7] P. M. Bakshi and K. T. Mahanthappa, "Expectation value formalism in quantum field theory, I," J. Math. Phys. **4**, 1 (1963).

[8] P. M. Bakshi and K. T. Mahanthappa, "Expectation value formalism in quantum field theory, II," J. Math. Phys. **4**, 12 (1963).

Chapter 4

In and Out states

Very often it happens that the Hamiltonian for a given problem has the property that as $t \to \pm\infty$,

$$H(t) = \begin{cases} H_{\rm out}\,, & \text{as } t \to +\infty, \\ H_{\rm in}\,, & \text{as } t \to -\infty. \end{cases} \tag{4.1}$$

For example, a parametric Hamiltonian of the form,

$$H(t) = \frac{1}{2}\left[\, P^2 + \omega^2(t)\, Q^2 \,\right], \qquad \text{where} \qquad \omega(t) = \begin{cases} \omega_{\rm out}\,, & \text{as } t \to +\infty, \\ \omega_{\rm in}\,, & \text{as } t \to -\infty. \end{cases} \tag{4.2}$$

is of this type. Here, the *in*-states oscillate at a frequency $\omega_{\rm in}$ whereas the *out*-states oscillate at a frequency $\omega_{\rm out}$. We know the eigenvalues and eigenstate of the initial and final system, but not the dynamics in between.

A second example is a parametric Hamiltonian of the form

$$H(t) = H_0 + H_1(t)\,, \qquad \text{with } H_1(t) \to 0 \text{ as } t \to \pm\infty. \tag{4.3}$$

Here the Hamiltonian is divided into two parts, one of which is time-independent and the other time-dependent. The time-dependent part vanishes as $t \to \pm\infty$. In this case $H_{\rm out} = H_{\rm in}$, but the Hamiltonian changes as a function of time. In both cases, we can define in and out states as solutions of the in and out Hamiltonian,

$$H_{\rm in}\,|\,\Psi_{\rm in}\,\rangle = E_{\rm in}\,|\,\Psi_{\rm in}\,\rangle\,, \qquad H_{\rm out}\,|\,\Psi_{\rm out}\,\rangle = E_{\rm out}\,|\,\Psi_{\rm out}\,\rangle\,. \tag{4.4}$$

Solutions of the time-dependent Hamiltonian are related to eigenvectors of the in and out Hamiltonians by the time-development operator,

$$|\,\Psi(t)\,\rangle = U(t,-\infty)\,|\,\Psi_{\rm in}\,\rangle = U(t,+\infty)\,|\,\Psi_{\rm out}\,\rangle\,, \tag{4.5}$$

so

$$|\,\Psi_{\rm out}\,\rangle = U^\dagger(t,+\infty)\,U(t,-\infty)\,|\,\Psi_{\rm in}\,\rangle = U(+\infty,-\infty)\,|\,\Psi_{\rm in}\,\rangle\,. \tag{4.6}$$

Here the out state depends on what the in state is. The transition amplitude $\mathcal{S}_{\rm out,in}$ for obtaining a *specific* out-state at $t = +\infty$ starting from a *specific* in state at $t = -\infty$ is given by

$$\mathcal{S}_{\rm out,in} = \langle\,\Psi_{\rm out}\,|\,U(+\infty,-\infty)\,|\,\Psi_{\rm in}\,\rangle\,, \tag{4.7}$$

the probability being the absolute magnitude of this amplitude. In this chapter, we develop methods to calculate $\mathcal{S}_{\rm out,in}$. For Hamiltonians of the form (4.3), if the matrix elements of $H_1(t)$ in eigenstates of H_0 are small, we can use perturbation theory to get an approximate answer. Here it is useful to introduce a new representation, called the **interaction** representation, in order to carry out the calculation. We discuss this new representation in the next section.

4.1 The interaction representation

So suppose that the Hamiltonian is of the form given in Eq. (4.3). Schrödinger's equation for this problem is:

$$[\,H_0 + H_1(t)\,]\,|\,\psi(t)\,\rangle = i\hbar\,\frac{\partial}{\partial t}|\,\psi(t)\,\rangle\,, \tag{4.8}$$

with H_0 and $H_1(t)$ Hermitian. Let H_0 satisfy the eigenvalue problem:

$$H_0\,|\,n\,\rangle = E_n\,|\,n\,\rangle\,, \qquad \text{with} \qquad \langle\,n\,|\,n'\,\rangle = \delta_{n,n'}\,. \tag{4.9}$$

We can remove the H_0 factor from the time-dependent part of the problem by setting:

$$|\,\psi(t)\,\rangle = e^{-iH_0t/\hbar}\,|\,\phi(t)\,\rangle\,. \tag{4.10}$$

Then $|\,\phi(t)\,\rangle$ satisfies:

$$H_1'(t)\,|\,\phi(t)\,\rangle = i\hbar\,\frac{\partial}{\partial t}|\,\phi(t)\,\rangle\,, \qquad \text{where} \qquad H_1'(t) = e^{+iH_0t/\hbar}\,H_1(t)\,e^{-iH_0t/\hbar}\,. \tag{4.11}$$

This representation of the dynamics is called the **interaction** representation. We can now put formally:

$$|\,\phi(t)\,\rangle = U'(t,t')\,|\,\phi(t')\,\rangle\,, \tag{4.12}$$

where $U'(t,t')$ is a time translation operator in the *interaction* representation. We will find an expression for $U'(t,t')$ as an expansion in powers of $H_1(t)$ in the next section. Initial and final conditions on the interaction representation state vector are:

$$|\,\phi(t)\,\rangle = \begin{cases} |\,n\,\rangle & \text{as } t \to +\infty, \\ |\,n'\,\rangle & \text{as } t \to -\infty, \end{cases} \tag{4.13}$$

both of which are eigenstates of H_0. From our discussion in the previous section, the transition amplitude for a transition from the in state $|\,n'\,\rangle$ to the out state $|\,n\,\rangle$ is given by

$$\mathcal{S}_{n,n'} = \langle\,n\,|\,U'(+\infty,-\infty)\,|\,n'\,\rangle\,. \tag{4.14}$$

4.2 The time development operator

When the Hamiltonian has an explicit dependence on time from some external source, the time development operator $U(t)$ no longer has the simple form:

$$U(t) = e^{-iHt/\hbar}\,, \tag{4.15}$$

and we must revisit the derivation of the operator. Such a situation occurred in the interaction representation discussed in the last section. When $H(Q,P,t)$ has an explicit time dependence, the time-development operator $U(t)$ satisfies Eq. (2.24), which we found in the last chapter:

$$H(Q,P,t)\,U(t,t') = i\hbar\frac{\partial U(t,t')}{\partial t}\,, \qquad \text{and} \qquad U^\dagger(t,t')\,H(Q,P,t) = -i\hbar\frac{\partial U^\dagger(t,t')}{\partial t}\,. \tag{4.16}$$

We have introduced an initial time variable t', and put $|\,\psi(t)\,\rangle = U(t,t')\,|\,\psi(t')\,\rangle$, with $U(t,t) = 1$. Here we are in the Schrödinger representation so that Q and P are time-independent — as a result, we will drop explicit reference to them in the following. Eqs. (4.16) can be written as integral equations of the form:

$$U(t,t') = 1 - \frac{i}{\hbar}\int_{t'}^{t} H(t_1)\,U(t_1,t')\,dt_1\,, \qquad U^\dagger(t,t') = 1 + \frac{i}{\hbar}\int_{t'}^{t} U^\dagger(t_1,t')\,H(t_1)\,dt_1\,. \tag{4.17}$$

Iterating the first of Eq. (4.17), we find:

$$U(t,t') = 1 + \left(\frac{-i}{\hbar}\right)\int_{t'}^{t} dt_1\, H(t_1) + \left(\frac{-i}{\hbar}\right)^2 \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2\, H(t_1)H(t_2) + \cdots .$$

Interchanging the order of integration in the last term gives:

$$\int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2\, H(t_1)H(t_2) = \int_{t'}^{t} dt_2 \int_{t_2}^{t} dt_1\, H(t_1)H(t_2) = \int_{t'}^{t} dt_1 \int_{t_1}^{t} dt_2\, H(t_2)H(t_1)\,.$$

So we find that the last term can be written:

$$\frac{1}{2}\int_{t'}^{t} dt_1 \int_{t'}^{t} dt_2\, [\Theta(t_1-t_2)H(t_1)H(t_2) + \Theta(t_2-t_1)H(t_2)H(t_1)] = \frac{1}{2}\int_{t'}^{t} dt_1 \int_{t'}^{t} dt_2\, \mathcal{T}\{\, H(t_1)H(t_2)\,\}$$

where T is the time-ordered product, defined by

$$\mathcal{T}\{\, H(t_1)H(t_2)\,\} = \Theta(t_1-t_2)H(t_1)H(t_2) + \Theta(t_2-t_1)H(t_2)H(t_1)\,,$$

and has the effect of time ordering the operators from right to left. Continuing in this way, we see that $U(t,t')$ has the expansion:

$$\begin{aligned} U(t,t') &= 1 + \sum_{n=1}^{\infty} \frac{1}{n!}\left(\frac{-i}{\hbar}\right)^n \int_{t'}^{t} dt_1 \int_{t'}^{t} dt_2 \cdots \int_{t'}^{t} dt_n \mathcal{T}\{\, H(t_1)H(t_2)\cdots H(t_n)\,\} \\ &= \mathcal{T}\Big\{\exp\Big[-\frac{i}{\hbar}\int_{t'}^{t} dt''\, H(Q,P,t'')\Big]\Big\}\,. \end{aligned} \tag{4.18}$$

In a similar way, iteration of the second of (4.17) gives:

$$U^\dagger(t,t') = 1 + \left(\frac{i}{\hbar}\right)\int_{t'}^{t} dt_1\, H(t_1) + \left(\frac{i}{\hbar}\right)^2 \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2\, H(t_2)H(t_1) + \cdots .$$

in this case, we find:

$$\begin{aligned} U^\dagger(t,t') &= 1 + \sum_{n=1}^{\infty} \frac{1}{n!}\left(\frac{i}{\hbar}\right)^n \int_{t'}^{t} dt_1 \int_{t'}^{t} dt_2 \cdots \int_{t'}^{t} dt_n \mathcal{T}^*\{\, H(t_1)H(t_2)\cdots H(t_n)\,\} \\ &= \mathcal{T}^*\Big\{\exp\Big[+\frac{i}{\hbar}\int_{t'}^{t} dt''\, H(Q,P,t'')\Big]\Big\}\,, \end{aligned} \tag{4.19}$$

where $\mathcal{T}^*$ is the anti-time-ordered product, defined by

$$\mathcal{T}^*\{\, H(t_1)H(t_2)\,\} = \Theta(t_2-t_1)H(t_1)H(t_2) + \Theta(t_1-t_2)H(t_2)H(t_1)\,.$$

which has the effect of time ordering from left to right, rather than right to left as in the time-ordered product. The Hamiltonian $H(Q,P,t)$ in Eqs. 4.18 and 4.19 is in the Schrödinger representation, with Q and P time-independent. We can also work out the time-development operator with the operators $Q(t)$ and $P(t)$ in the Heisenberg representation. For this case, we start with Eqs. (2.24):

$$i\hbar\,\frac{\partial U(t,t')}{\partial t} = U(t,t')\,H(Q(t),P(t),t)\,, \qquad \text{and} \qquad -i\hbar\,\frac{\partial U^\dagger(t,t')}{\partial t} = H(Q(t),P(t),t)\,U^\dagger(t,t')\,, \tag{4.20}$$

which can be written as the following integral equations:

$$U(t,t') = 1 - \frac{i}{\hbar}\int_{t'}^{t} U(t_1,t')\,H(Q(t_1),P(t_1),t_1)\,\mathrm{d}t_1\,, \tag{4.21}$$

$$U^\dagger(t,t') = 1 + \frac{i}{\hbar}\int_{t'}^{t} H(Q(t_1),P(t_1),t_1)\,U^\dagger(t_1,t')\,\mathrm{d}t_1\,. \tag{4.22}$$

Iteration of these equations leads to results similar to what we found in the Schrödinger picture. We get in this case:

$$U(t,t') = \mathcal{T}^*\Big\{\exp\Big[-\frac{i}{\hbar}\int_{t'}^{t} dt''\, H(Q(t''),P(t''),t'')\Big]\Big\}\,, \tag{4.23}$$

$$U^\dagger(t,t') = \mathcal{T}\ \Big\{\exp\Big[+\frac{i}{\hbar}\int_{t'}^{t} dt''\, H(Q(t''),P(t''),t'')\Big]\Big\}\,. \tag{4.24}$$

We note here that $U^\dagger(t,t') = U(t',t)$. In addition, one can show that $U(t_1,t_2)\,U(t_2,t_3) = U(t_1,t_3)$.

4.3 Forced oscillator

As an example of the use of the interaction representation and the perturbation expansion of the time-development operator, we study a forced harmonic oscillator with the Hamiltonian (See Chapter 16, Section 16.6 where we solved this problem exactly.):

$$H(t) = H_0 + H_1(t)\,, \tag{4.25}$$

$$H_0 = \frac{P^2}{2m} + \frac{1}{2}m\,\omega_0^2\,Q^2\,, \qquad \text{and} \qquad H_1(t) = -Q\,F(t)\,. \tag{4.26}$$

where $F(t)$ is an external force which commutes with Q and P, and where $F(t)\to 0$ as $t\to\pm\infty$. We first need to find the eigenvalues and eigenvectors for H_0. So we put:

$$Q = \sqrt{\frac{\hbar}{2m\,\omega_0}}\,\big(A + A^\dagger\big)\,, \qquad P = \sqrt{\frac{\hbar\,m\,\omega_0}{2}}\,\frac{1}{i}\,\big(A - A^\dagger\big)\,, \tag{4.27}$$

$$[\,Q,P\,] = i\hbar\,, \qquad [\,A,A^\dagger\,] = 1\,. \tag{4.28}$$

Then

$$H_0 = \frac{P^2}{2m} + \frac{1}{2}m\,\omega_0^2\,Q^2 = \hbar\omega_0\,\big[\,A^\dagger A + 1/2\,\big]\,, \qquad [\,H_0,A\,] = -\hbar\omega_0\,A\,. \tag{4.29}$$

The eigenvalues and eigenvectors are given by:

$$H_0\,|\,n\,\rangle = \hbar\omega_0\,\big[\,n + 1/2\,\big]\,|\,n\,\rangle\,, \qquad |\,n\,\rangle = \frac{(\,A^\dagger\,)^n}{\sqrt{n!}}\,|\,0\,\rangle\,. \tag{4.30}$$

Next, we want to find $H_1'(t)$ in the interaction representation. This is given by:

$$H_1'(t) = e^{+iH_0t/\hbar}\,H_1(t)\,e^{-iH_0t/\hbar} = -e^{+iH_0t/\hbar}\,Q\,e^{-iH_0t/\hbar}\,J(t) = -Q'(t)\,F(t)\,. \tag{4.31}$$

where

$$Q'(t) = e^{+iH_0t/\hbar}\,Q\,e^{-iH_0t/\hbar} = \sqrt{\frac{\hbar}{2m\omega_0}}\,\big(\,A'(t) + A'^\dagger(t)\,\big)\,, \tag{4.32}$$

with:

$$\begin{aligned} A'(t) &= e^{+iH_0t/\hbar}\,A\,e^{-iH_0t/\hbar} \\ &= A + \left(\frac{it}{\hbar}\right)\,[\,H_0,A\,] + \frac{1}{2!}\left(\frac{it}{\hbar}\right)^2\,[\,H_0,[\,H_0,A\,]\,] + \cdots \\ &= A + (-i\omega_0 t)\;A + \frac{1}{2!}\,(-i\omega_0 t)^2\;A + \cdots = e^{-i\omega_0 t}\,A\,. \end{aligned} \tag{4.33}$$

Here we have used Eq. (B.14) in Appendix **??**. Similarly,

$$A'^\dagger(t) = e^{+i\omega_0 t}\,A^\dagger\,. \tag{4.34}$$

So

$$Q'(t) = \sqrt{\frac{\hbar}{2m\omega_0}}\left(A\,e^{-i\omega_0 t} + A^\dagger\,e^{+i\omega_0 t}\right), \tag{4.35}$$

which satisfies the oscillator equation of motion:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega_0^2\right\} Q'(t) = 0\,. \tag{4.36}$$

So from (4.14), the probability of finding an out state $|\,n'\,\rangle$ of the free oscillator from an in state $|\,n'\,\rangle$ is given by:

$$P_{n,n'} = |\,{}_{\rm out}\langle\,n\,|\,U(+\infty,-\infty)\,|\,n'\,\rangle_{\rm in}\,|^2 \tag{4.37}$$

where the matrix element is given by (4.18):

$$\begin{aligned}
&{}_{\rm out}\langle\,n\,|\,U(+\infty,-\infty)\,|\,n'\,\rangle_{\rm in}\\
&\quad= \langle\,n\,|\left\{1 + \sum_{m=1}^{\infty}\frac{1}{m!}\left(\frac{-i}{\hbar}\right)^m\int_{-\infty}^{+\infty}dt_1\int_{-\infty}^{+\infty}dt_2\cdots\int_{-\infty}^{+\infty}dt_m\mathcal{T}\{\,H(t_1)H(t_2)\cdots H(t_m)\,\}\right\}|\,n'\,\rangle\\
&\quad= \delta_{n,n'} + \sum_{m=1}^{\infty}\frac{1}{m!}\left(\frac{i}{\hbar}\right)^m\int_{-\infty}^{+\infty}dt_1\int_{-\infty}^{+\infty}dt_2\cdots\int_{-\infty}^{+\infty}dt_m\,\tau^{(m)}_{n,n'}(t_1,t_2,\ldots,t_m)\,F(t_1)F(t_2)\cdots F(t_m)\,,
\end{aligned} \tag{4.38}$$

where

$$\tau^{(m)}_{n,n'}(t_1,t_2,\ldots,t_m) = \langle\,n\,|\,\mathcal{T}\{\,Q'(t_1)Q'(t_2)\cdots Q'(t_m)\,\}\,|\,n'\,\rangle\,. \tag{4.39}$$

These τ-functions are the time-ordered product of the position operator $Q'(t)$ in the interaction representation, and are the quantities we want to calculate here. Let us first look at the τ-functions for the $n'=0$ to $n=0$ transition. That is, the amplitude that nothing happens to the oscillator after the application of the external force. For this case, we see that there are no odd-m terms, since we will always have a creation or destruction operator left over. So the first non-zero term is $\tau^{(2)}_{0,0}(t,t')$ which is easily calculated to be:

$$\tau^{(2)}_{0,0}(t,t') = \langle\,0\,|\,\mathcal{T}\{\,Q'(t)Q'(t')\,\}\,|\,0\,\rangle = \frac{\hbar}{2m\omega_0}\left\{e^{-i\omega_0(t-t')}\,\theta(t-t') + e^{i\omega_0(t-t')}\,\theta(t-t')\right\}, \tag{4.40}$$

where we have used (4.35). Comparing this to our result for $G_F(t-t')$ in Chapter 16, Example 42 on page 192, we see that $\tau^{(2)}_{0,0}(t,t')$ is proportional to $G_F(t-t')$, and we find:

$$\begin{aligned}
G_F(t-t') &= \frac{im}{\hbar}\,\tau^{(2)}_{0,0}(t,t') = \frac{im}{\hbar}\,\langle\,0\,|\,\mathcal{T}\{\,Q'(t)Q'(t')\,\}\,|\,0\,\rangle\\
&= \frac{i}{2\omega_0}\left\{e^{-i\omega_0(t-t')}\,\theta(t-t') + e^{i\omega_0(t-t')}\,\theta(t-t')\right\}.
\end{aligned} \tag{4.41}$$

That is, the Feynman Green function can be defined in terms of the time-ordered product of two positions operators. Let us make sure that the Feynman Green function defined in this way satisfies the correct differential equation. The first derivative is given by:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,G_F(t-t') = \frac{im}{\hbar}\,\langle\,0\,|\left(\dot{Q}'(t)Q'(t')\,\theta(t-t') + Q'(t')\dot{Q}'(t)\,\theta(t'-t)\right)|\,0\,\rangle\,, \tag{4.42}$$

and the second derivative is:

$$\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}t^2}\,G_F(t-t') &= \frac{im}{\hbar}\,\langle\,0\,|\left(\ddot{Q}'(t)Q'(t')\,\theta(t-t') + Q'(t')\ddot{Q}'(t)\,\theta(t'-t)\right)|\,0\,\rangle\\
&\qquad + \frac{i}{\hbar}\,\langle\,0\,|\left(P'(t)Q'(t) - Q'(t)P'(t)\right)|\,0\,\rangle\,\delta(t-t')\\
&= -\omega_0^2\,G_F(t-t') + \delta(t-t')\,,
\end{aligned} \tag{4.43}$$

where we have used the equations of motion given in Eq. (4.36). So $G_F(t-t')$ satisfies the required differential equation:

$$\left\{ \frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega_0^2 \right\} G_F(t-t') = \delta(t-t') \,. \tag{4.44}$$

We solved (4.44) in example 42 using Fourier transforms and contour integration, and found:

$$G_F(t-t') = -\int_F \frac{\mathrm{d}\omega}{2\pi} \frac{e^{-i\omega(t-t')}}{\omega^2-\omega_0^2} = \frac{i}{2\omega_0} \left\{ e^{-i\omega_0(t-t')}\,\theta(t-t') + e^{i\omega_0(t-t')}\,\theta(t'-t) \right\} , \tag{4.45}$$

in agreement with Eq. (4.41). So from Eq. (4.38), we find that the $m=2$ term in the expansion of the probability amplitude is given by::

$$\begin{aligned}
\frac{1}{2}\left(\frac{i}{\hbar}\right)^2 \int_{-\infty}^{+\infty} dt_1 \int_{-\infty}^{\infty} dt_2\, \tau^{(2)}_{0,0}(t_1,t_2)\, F(t_1)F(t_2) \\
&= \frac{i}{2\hbar m} \int_{-\infty}^{+\infty} dt_1 \int_{-\infty}^{\infty} dt_2\, G_F(t_1-t_2)\, F(t_1)\, F(t_2) \\
&= -\frac{i}{2\hbar m} \int_{-\infty}^{+\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \int_F \frac{\mathrm{d}\omega}{2\pi} \frac{e^{-i\omega(t_1-t_2)}}{\omega^2-\omega_0^2}\, F(t_1)\, F(t_2) \\
&= -\frac{i}{2\hbar m} \int_F \frac{\mathrm{d}\omega}{2\pi} \frac{|\,\tilde F(\omega)\,|^2}{(\omega-\omega_0)(\omega+\omega_0)} = -\frac{1}{2}\,|\,a\,|^2 \,, \qquad \text{where} \qquad |\,a\,|^2 = \frac{|\tilde F(\omega_0)|^2}{2\hbar\omega_0 m} \,.
\end{aligned} \tag{4.46}$$

Here we have assumed that $|\,\tilde F(\omega)\,|^2 \to 0$ as $\omega \to \pm i\infty$ so that we can close the contour in either the UHP or the LHP.

The next term in the expansion (4.38) is the $m=4$ term. This requires that we calculate the ground state to ground state time-ordered product $\tau^{(4)}_{0,0}(t_1,t_2,t_3,t_4)$ given by:

$$\tau^{(4)}_{0,0}(t_1,t_2,t_3,t_4) = \langle\, 0\,|\, \mathcal{T}\{\, Q'(t_1)Q'(t_2)Q'(t_3)Q'(t_4)\,\}\,|\,0\,\rangle \,. \tag{4.47}$$

There is a theorem we can use in this case, called **Wick's theorem**. In order to state Wick's theorem, we need the following definition of the **normal ordered product**.

Definition 10 (Normal ordered product)**.** The normal ordered product, denoted by:

$$\mathcal{N}\{\, Q(t_1)Q(t_2)\cdots Q(t_m)\,\} \,, \tag{4.48}$$

is defined to be the product of all operators such that all creation operators stand to the left and all annihilation operators stand on the right. For example for two operators, we have:

$$\mathcal{N}\{\, Q(t_1)Q(t_2)\,\} = \frac{\hbar}{2\omega_0 m} \left\{ AA\, e^{-i\omega_0(t_1+t_2)} + A^\dagger A\, e^{i\omega_0(t_1-t_2)} + A^\dagger A\, e^{i\omega_0(t_2-t_1)} + A^\dagger A^\dagger\, e^{i\omega_0(t_1+t_2)} \right\} . \tag{4.49}$$

By construction, we have that:

$$\langle\, 0\,|\mathcal{N}\{\, Q(t_1)Q(t_2)\cdots Q(t_m)\,\}|\,0\,\rangle = 0 \,. \tag{4.50}$$

Theorem 8 (Wick's theorem)**.** *The time-ordered product of m operators can be expanded by the expression:*

$$\begin{aligned}
\mathcal{T}\{\, Q(t_1)Q(t_2)\cdots Q(t_m)\,\} &= \mathcal{N}\{\, Q(t_1)Q(t_2)\cdots Q(t_m)\,\} \\
&\quad + \sum_{perm} \langle\, 0\,|\mathcal{T}\{\, Q(t_1)Q(t_2)\,\}|\,0\,\rangle\, \mathcal{N}\{\, Q(t_3)\cdots Q(t_m)\,\} \\
&\quad + \sum_{perm} \langle\, 0\,|\mathcal{T}\{\, Q(t_1)Q(t_2)\,\}|\,0\,\rangle\, \langle\, 0\,|\mathcal{T}\{\, Q(t_2)Q(t_3)\,\}|\,0\,\rangle\, \mathcal{N}\{\, Q(t_4)\cdots Q(t_m)\,\} \\
&\quad + \cdots + \sum_{perm} \langle\, 0\,|\mathcal{T}\{\, Q(t_1)Q(t_2)\,\}|\,0\,\rangle\, \langle\, 0\,|\mathcal{T}\{\, Q(t_2)Q(t_3)\,\}|\,0\,\rangle\, \cdots \langle\, 0\,|\mathcal{T}\{\, Q(t_{m-1})Q(t_m)\,\}|\,0\,\rangle \,,
\end{aligned} \tag{4.51}$$

for m even. For m odd, the last line is replaced by:

$$\sum_{perm} \langle\, 0\,|\mathcal{T}\{\, Q(t_1)Q(t_2)\,\}|\,0\,\rangle\,\langle\, 0\,|\mathcal{T}\{\, Q(t_2)Q(t_3)\,\}|\,0\,\rangle\,\cdots\langle\, 0\,|\mathcal{T}\{\, Q(t_{m-2})Q(t_{m-1})\,\}|\,0\,\rangle\, Q(t_m)\,. \tag{4.52}$$

Here $\mathcal{N}\{\,\cdot\,\}$ denotes the normal ordered product of operators, defined above.

Proof. The theorem is proved by induction in several books (see, for example, Drell [1]), and will not be reproduced here. One can see that to move from a time-ordered product to a normal ordered product, one must commute creation and annihilation operators which give c-numbers. These c-numbers can then be factored out and thus reducing the number of quantities to normal order by two. One continues in this way until all terms are commuted into normal ordered operators. □

So for our case, application of Wick's theorem to the case of $m = 4$ gives:

$$\begin{aligned}\tau^{(4)}_{0,0}(t_1,t_2,t_3,t_4) &= \langle\, 0\,|\,\mathcal{T}\{\, Q'(t_1)Q'(t_2)Q'(t_3)Q'(t_4)\,\}\,|\,0\,\rangle \\ &= \tau^{(2)}_{0,0}(t_1,t_2)\,\tau^{(2)}_{0,0}(t_3,t_4) + \tau^{(2)}_{0,0}(t_1,t_3)\,\tau^{(2)}_{0,0}(t_2,t_4) + \tau^{(2)}_{0,0}(t_1,t_4)\,\tau^{(2)}_{0,0}(t_2,t_3)\,.\end{aligned} \tag{4.53}$$

So from (4.38), the $m = 4$ term contributes a factor:

$$\frac{1}{4!}\left(\frac{i}{\hbar}\right)^4 \int_{-\infty}^{+\infty} dt_1 \int_{-\infty}^{+\infty} dt_2 \int_{-\infty}^{+\infty} dt_3 \int_{-\infty}^{+\infty} dt_4\, \tau^{(m)}_{0,0}(t_1,t_2,t_3,t_4)\, F(t_1)F(t_2)F(t_3)F(t_4) = \frac{3}{4!}\,|a|^4\,. \tag{4.54}$$

So from Eq. (4.38), we find to 4^{th} order:

$$\langle\, 0\,|\,U(+\infty,-\infty)\,|\,0\,\rangle = 1 - \frac{|\,a\,|^2}{2} + \frac{1}{2!}\left(\frac{|\,a\,|^2}{2}\right)^2 + \cdots \approx e^{-|a|^2/2}\,, \tag{4.55}$$

so that from (4.37):

$$P_{0,0} = |\,{}_{\text{out}}\langle\, 0\,|\,U(+\infty,-\infty)\,|\,0\,\rangle_{\text{in}}\,|^2 = 1 - |a|^2 + \frac{1}{2!}\,|a|^4 \cdots \approx e^{-|a|^2} \tag{4.56}$$

in agreement with the exact result found in Chapter 16, Eq. (16.129) on page 191 to second order in $|a|^2$. In this section, we only computed ground state to ground state probabilities. In order to compute ground state ($n' = 0$) to excited state (n) probabilities requires similar calculations, except now the only terms that contribute are those that have n factors of $Q'(t)$ left over from the normal ordered product. With the help of Wick's theorem, the first order term can be easily calculated. The next order, however, is hard to do this way, and we will learn other methods in the chapter on path integrals.

Exercise 7. Use Wick's theorem to find the first-order contribution to the probability $P_{4,0}$ for the forced harmonic oscillator. Show your answer agrees with the exact result.

References

[1] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, NY, 1964).

Chapter 5

Density matrix formalism

5.1 Classical theory

In classical systems, we often want to solve a problem where we are given a distribution $\rho_0(q_0, p_0)$ of values of q and p at $t = 0$, and we want to find the distribution $\rho(q, p, t)$ at time t later, where the values of $q(t)$ and $p(t)$ have evolved according to the classical equations of motion,

$$\dot{q} = \{\, q, H(q,p)\,\} = \frac{\partial H(q,p)}{\partial p}\,, \qquad \dot{p} = \{\, p, H(q,p)\,\} = -\frac{\partial H(q,p)}{\partial q}\,, \tag{5.1}$$

with initial conditions, $q(0) = q_0$ and $p(0) = p_0$. In this section, we consider only Hamiltonians which do not depend *explicitly* on time. If we don't *loose* any trajectories of $q(t)$ and $p(t)$ in phase space from the initial distribution, the equation of motion for $\rho(q, p, t)$ is given by

$$\frac{\mathrm{d}\rho(q,p,t)}{\mathrm{d}t} = \frac{\partial\rho(q,p,t)}{\partial t} + \dot{q}\,\frac{\partial\rho(q,p,t)}{\partial q} + \dot{p}\,\frac{\partial\rho(q,p,t)}{\partial p} = \frac{\partial\rho(q,p,t)}{\partial t} + \{\,\rho(q,p,t), H(q,p)\,\} = 0\,, \tag{5.2}$$

with initial conditions $\rho(q, p, 0) = \rho_0(q, p)$. For Hamiltonians of the form $H(q, p, t) = p^2/2m + V(q)$, Eq. (5.2) can be written as

$$\frac{\partial\rho(q,p,t)}{\partial t} + \frac{p}{m}\,\frac{\partial\rho(q,p,t)}{\partial q} + F(q)\,\frac{\partial\rho(q,p,t)}{\partial p} = 0\,, \tag{5.3}$$

where $F(q) = -\partial V(q)/\partial q$, which is called Boltzmann's equation. The time evolution of the coordinates in terms of the initial coordinates are given by equations of the form, $q(q_0, p_0, t)$ and $p(q_0, p_0, t)$. Both the set (q, p) at time t and the set (q_0, p_0) at time $t = 0$ are canonical coordinates and can be used to find Poisson equations of motion. In particular, the Hamiltonian is a constant, independent of time,

$$\frac{\mathrm{d}H(q,p)}{\mathrm{d}t} = \frac{\partial H(q,p)}{\partial t} = 0\,, \tag{5.4}$$

so $H[\,q(t), p(t)\,] = H(q_0, p_0)$. We turn next to iterative solutions of these equations.

5.1.1 Classical time development operator

In this section, we consider only Hamiltonians which do not depend *explicitly* on time. Then We first note that we can find iterative solutions to the Poisson bracket equations of motion by writing them as integral

equations. Integrating (5.1) over t from $t = 0$ to t and iterating over and over again gives

$$\begin{aligned} q(t) &= q_0 + \int_0^t \mathrm{d}t' \, \{ \, q(t'), H(q,p) \, \} \\ &= q_0 + \int_0^t \mathrm{d}t' \, \{ \, q_0, H(q_0,p_0) \, \} + \int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}t'' \, \{ \, \{ \, q_0, H(q_0,p_0) \, \}, H(q_0,p_0) \, \} + \cdots \, , \\ &= q_0 + t \, \{ \, q_0, H(q_0,p_0) \, \} + \frac{t^2}{2!} \, \{ \, \{ \, q_0, H(q_0,p_0) \, \}, H(q_0,p_0) \, \} + \cdots \, , \end{aligned} \tag{5.5}$$

with a similar expression for $p(t)$. Here it is simplest to compute the Poisson brackets with respect to the set (q_0, p_0). The equation of motion (5.2) for the distributions function $\rho(q,p,t)$ can be written as

$$\frac{\partial \rho(q,p,t)}{\partial t} = -\{ \, \rho(q,p,t), H(q,p,t) \, \} \, , \tag{5.6}$$

which has an opposite sign from the equations of motion in Eqs. (5.1), so the iterated solution for $\rho(q,p,t)$ becomes

$$\rho(q,p,t) = \rho_0(q,p) - t \, \{ \, \rho_0(q,p), H(q,p) \, \} + \frac{t^2}{2!} \, \{ \, \{ \, \rho_0(q,p), H(q,p) \, \}, H(q,p) \, \} + \cdots \, . \tag{5.7}$$

Here we compute the Poisson brackets with respect to the *final* set (q,p). From expressions (5.5) and (5.7), we see that we can define an operator for the time-evoution of the system. Let us define a classical time-development operator by

Definition 11 (classical time-development operator)**.** The time-development operator $U_{\rm op}(t)$ is a *right-action* operator, defined by:

$$U_{\rm op}(t) = e^{t \, H_{\rm op}} \, , \qquad \text{where} \qquad H_{\rm op} := \{ \quad , H \} = \frac{\partial H(q_0,p_0)}{\partial p_0} \, \frac{\partial}{\partial q_0} - \frac{\partial H(q_0,p_0)}{\partial q_0} \, \frac{\partial}{\partial p_0} \, , \tag{5.8}$$

where the slot in the expression $\{ \quad , H\}$ is the position where we put the quantity which is to be operated on. We can just as well compute the Poisson brackets with respect to the set (q,p) at time t (see below). Note that $t \, H_{\rm op}$ is dimensionless.

So with this definition, we find

$$q(t) = U_{\rm op}(t) \, q_0 \, , \qquad p(t) = U_{\rm op}(t) \, p_0 \, , \tag{5.9}$$

and in general

$$A(q(t),p(t)) = U_{\rm op}(t) \, A(q_0,p_0) \, . \tag{5.10}$$

Also, since

$$H_{\rm op} \, A(q_0,p_0) = \{ \, A(q_0,p_0), H(q_0,p_0) \, \} \, ,$$

we find that

$$H_{\rm op}^2 \, A(q_0,p_0) = H_{\rm op} \, \{ \, A(q_0,p_0), H(q_0,p_0) \, \} = \{ \, \{ \, A(q_0,p_0), H(q_0,p_0) \, \}, H(q_0,p_0) \, \} \, .$$

We will need the following theorem.

Theorem 9.

$$U_{op}(t_1) \, U_{op}(t_2) = U_{op}(t_1 + t_2) \, . \tag{5.11}$$

Proof. We have:

$$\begin{aligned} U_{\rm op}(t_1)\,U_{\rm op}(t_2)\,A(q_0,p_0) &= e^{t_1\,H_{\rm op}}\,\left[e^{t_2\,H_{\rm op}}\,A(q_0,p_0)\right] \\ &= \sum_{n=0}^{\infty}\sum_{m=1}^{\infty}\frac{t_1^n\,t_2^m}{n!\,m!}H_{\rm op}^n\,[\,H_{\rm op}^m\,A(q_0,p_0)\,] = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{t_1^n\,t_2^m}{n!\,m!}\,[\,H_{\rm op}^{m+n}\,A(q_0,p_0)\,] \\ &= \sum_{k=0}^{\infty}[\,H_{\rm op}^k\,A(q_0,p_0)\,]\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\delta_{k,n+m}\,\frac{t_1^n\,t_2^m}{n!\,m!} = \sum_{k=0}^{\infty}\frac{(t_1+t_2)^k}{k!}\,[\,H_{\rm op}^k\,A(q_0,p_0)\,] \\ &= e^{(t_1+t_2)\,H_{\rm op}}\,A(q_0,p_0) \;= U_{\rm op}(t_1+t_2)\,A(q_0,p_0)\,, \end{aligned}$$

where we have used the binomial theorem. □

Using this theorem, we see that since $U_{\rm op}(0) = 1$, we have $U_{\rm op}(t)\,U_{\rm op}(-t) = 1$ so that the inverse operator is given by $U_{\rm op}^{-1}(t) = U_{\rm op}(-t)$. This means that the *operator*, $U_{\rm op}(t)A(q_0,p_0)U_{\rm op}(-t)$ has no effect when operating on any function $B(q_0,p_0)$. We state this in the form of another curious theorem as follows:

Theorem 10.

$$U_{op}(t)\,A(q_0,p_0)\,U_{op}^{-1}(t)\,B(q_0,p_0) = A[\,q(t),p(t)\,]\,B(q_0,p_0)\,. \tag{5.12}$$

Proof. We find

$$\begin{aligned} U_{\rm op}(t)\,A(q_0,p_0)\,U_{\rm op}^{-1}(t)\,B(q_0,p_0) &= U_{\rm op}(t)\,A(q_0,p_0)\,[\,U_{\rm op}(-t)\,B(q_0,p_0)\,] \\ &= U_{\rm op}(t)\,[\,A(q_0,p_0)\,B[\,q(-t),p(-t)\,]\,] = A[\,q(t),p(t)\,]\,B(q_0,p_0)\,. \end{aligned}$$

□

So the operator

$$A_{\rm op}[\,q(t),p(t)\,] = U_{\rm op}(t)\,A(q_0,p_0)\,U_{\rm op}^{-1}(t)\,, \tag{5.13}$$

changes only $A(q_0,p_0)$, and does nothing to any function $B(q_0,p_0)$ to the right of this operator. In particular, we can write the *operator* relations

$$\begin{aligned} q_{\rm op}(t) &= U_{\rm op}(t)\,q_0\,U_{\rm op}^{-1}(t)\,, \\ p_{\rm op}(t) &= U_{\rm op}(t)\,p_0\,U_{\rm op}^{-1}(t)\,, \end{aligned} \tag{5.14}$$

which have the *value* $q(t)$ and $p(t)$ when operating on any function $f(q_0,p_0)$,

$$\begin{aligned} q_{\rm op}(t)\,f(q_0,p_0) &= q(t)\,f(q_0,p_0)\,, \\ p_{\rm op}(t)\,f(q_0,p_0) &= p(t)\,f(q_0,p_0)\,. \end{aligned} \tag{5.15}$$

For the density function $\rho(q,p,t)$, we can use the same time evolution operator if we compute the Poisson brackets with respect to the final set of coordinates (q,p) and with an opposite sign. That is Eq. (5.7) can be written as

$$\rho(q,p,t) = U_{\rm op}^{-1}(t)\,\rho_0(q,p)\,. \tag{5.16}$$

Here we regard (q,p) as dummy variables. Again, we can define a density *operator* $\rho_{\rm op}(q,p,t)$ by the expression

$$\rho_{\rm op}(q,p,t) = U_{\rm op}^{-1}(t)\,\rho_0(q,p)\,U_{\rm op}(t)\,, \tag{5.17}$$

which has the value $\rho(q,p,t)$ when operating on any function $f(q,p)$,

$$\rho_{\rm op}(q,p,t)\,f(q,p) = \rho(q,p,t)\,f(q,p)\,. \tag{5.18}$$

5.1.2 Classical averages

We interpret $\rho(q,p,t)$ as the *probability* of finding the system at a point (q,p) in phase space at time t. The next theorem expresses conservation of this probability, a result known as Liouville's theorem.

Theorem 11. *The distribution function $\rho(q,p,t)$ is normalized according to*

$$\mathcal{N}(t) = \iint_{-\infty}^{+\infty} \rho(q,p,t)\,\frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar} = 1\,, \tag{5.19}$$

for all t.

Proof. Differentiating (5.19) with respect to t and using (5.6) gives

$$\begin{aligned}\frac{\mathrm{d}\mathcal{N}(t)}{\mathrm{d}t} &= -\iint_{-\infty}^{+\infty} \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\left\{\frac{\partial\rho(q,p,t)}{\partial q}\,\frac{\partial H(q,p,t)}{\partial p} - \frac{\partial\rho(q,p,t)}{\partial p}\,\frac{\partial H(q,p,t)}{\partial q}\right\} \\ &= \iint_{-\infty}^{+\infty} \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\rho(q,p,t)\left\{\frac{\partial^2 H(q,p,t)}{\partial q\,\partial p} - \frac{\partial^2 H(q,p,t)}{\partial p\,\partial q}\right\} = 0\,.\end{aligned} \tag{5.20}$$

Here we have integrated by parts and assumed that $\rho(q,p,t) \to 0$ as either q or p go to $\pm\infty$. □

So we conclude that

$$\iint_{-\infty}^{+\infty} \rho(q,p,t)\,\frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar} = \iint_{-\infty}^{+\infty} \rho_0(q_0,p_0)\,\frac{\mathrm{d}q_0\,\mathrm{d}p_0}{2\pi\hbar} = 1\,. \tag{5.21}$$

The average value of q or p at any time t can be computed in two ways: we can either solve Eq. (5.6) for $\rho(q,p,t)$ with initial value $\rho_0(q,p)$, and average over q and p, or solve Eq. (5.1) for $q(t)$ and $p(t)$ with initial values of q_0 and p_0, and average over the initial values $\rho_0(q_0,p_0)$. We state this in the form of a theorem.

Theorem 12.

$$\langle\, q(t)\,\rangle = \iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,q\,\rho(q,p,t) = \iint \frac{\mathrm{d}q_0\,\mathrm{d}p_0}{2\pi\hbar}\,q(\,q_0,p_0,t\,)\,\rho_0(q_0,p_0)\,, \tag{5.22a}$$

$$\langle\, p(t)\,\rangle = \iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,p\,\rho(q,p,t) = \iint \frac{\mathrm{d}q_0\,\mathrm{d}p_0}{2\pi\hbar}\,p(\,q_0,p_0,t\,)\,\rho_0(q_0,p_0)\,. \tag{5.22b}$$

Remark 12. This first method of averaging corresponds in quantum mechanics to computing averages in the Schrödinger representation, whereas the second method corresponds to computing averages in the Heisenberg representation.

Example 24. We show directly that time time derivative of the average field is the same using both method of averaging. Using the *first* method of averaging, we find

$$\begin{aligned}\frac{\partial\langle\, q(t)\,\rangle}{\partial t} &= \iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\frac{\partial\rho(q,p,t)}{\partial t}\,q = -\iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\{\,\rho(q,p,t),H\,\}\,q \\ &= -\iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\left\{\frac{\partial\rho(q,p,t)}{\partial q}\,\frac{\partial H}{\partial p} - \frac{\partial\rho(q,p,t)}{\partial p}\,\frac{\partial H}{\partial q}\right\}q \\ &= \iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\rho(q,p,t)\left\{\frac{\partial}{\partial q}\Big[\,q\,\frac{\partial H}{\partial p}\,\Big] - \frac{\partial}{\partial p}\Big[\,q\,\frac{\partial H}{\partial q}\,\Big]\right\} \\ &= \iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\rho(q,p,t)\,\{\,q,H\,\} = \iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\rho(q,p,t)\,\dot{q}\,.\end{aligned} \tag{5.23}$$

Since q and p are dummy integration variables, this agrees with computing the time derivative using the *second* method.

5.1.3 Classical correlation and Green functions

We start by introducing a number of definitions of classical correlation and Green functions, which will be useful later.

Definition 12 (correlation coefficient)**.** The correlation coefficient $F(t,t')$ is defined by

$$F(t,t') = \langle\, q(t)\, q(t')\,\rangle = \iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\rho(q,p)\,q(t)\,q(t')\,. \tag{5.24}$$

Definition 13 (spectral function)**.** The spectral function $\sigma(t,t')$ is defined as the expectation value of the Poisson bracket of $q(t)$ and $q(t')$ by

$$\sigma(t,t') = \langle\, \{\, q(t), q(t')\,\}\,\rangle = \iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\rho(q,p)\,\{\, q(t), q(t')\,\}\,. \tag{5.25}$$

Definition 14 (Green functions)**.** Advanced and retarded Green functions are defined by

$$G_{\mathrm{A}}(t,t') = +\sigma(t,t')\,\Theta(t'-t)\,, \tag{5.26a}$$
$$G_{\mathrm{R}}(t,t') = -\sigma(t,t')\,\Theta(t-t')\,. \tag{5.26b}$$

It will be useful to introduce a matrix of correlation and Green functions as follows. We first define the matrix $\overline{\mathbb{G}}(t,t')$ of Green functions by

$$\overline{\mathbb{G}}(t,t') = \begin{pmatrix} 2i\,F(t,t') & G_{\mathrm{A}}(t,t') \\ G_{\mathrm{R}}(t,t') & 0 \end{pmatrix}\,. \tag{5.27}$$

We will find it useful later to define new Green functions by a change of basis of this matrix with the following definition. We define $\mathbb{G}(t,t')$ by

$$\begin{aligned} \mathbb{G}(t,t') = \mathbb{U}\,\overline{\mathbb{G}}(t,t')\,\mathbb{U}^{-1} &= \begin{pmatrix} G_{++}(t,t') & G_{+-}(t,t') \\ G_{-+}(t,t') & G_{--}(t,t') \end{pmatrix}\,, \\ &= \mathbb{H}(t,t')\,G_{>}(t,t') + \mathbb{H}^{T}(t',t)\,G_{<}(t,t')\,, \end{aligned} \tag{5.28}$$

where

$$\mathbb{U} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\,, \qquad \mathbb{U}^{-1} = \mathbb{U}^{T} = \mathbb{U}^{\dagger} = \mathbb{U} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\,. \tag{5.29}$$

Multiplying this out, we find:

$$\begin{aligned} G_{++}(t,t') &= \Theta(t-t')\,G_{>}(t,t') + \Theta(t'-t)\,G_{<}(t,t')\,, \\ G_{-+}(t,t') &= G_{>}(t,t') \\ G_{+-}(t,t') &= G_{<}(t,t') \\ G_{--}(t,t') &= \Theta(t'-t)\,G_{>}(t,t') + \Theta(t-t')\,G_{<}(t,t')\,, \end{aligned} \tag{5.30}$$

where

$$G_{\gtrless}(t,t')/i = F(t,t') \pm i\sigma(t,t')/2 = \mathrm{Tr}\big[\,[q(t)q(t') \pm i\{\, q(t), q(t')\,\}/2]\,\big]\,. \tag{5.31}$$

In (5.28), we have defined the matrices $\mathbb{H}(t,t')$ and its transpose $\mathbb{H}^{T}(t,t')$ by:

$$\mathbb{H}(t,t') = \begin{pmatrix} \Theta(t-t') & 0 \\ 1 & \Theta(t'-t) \end{pmatrix}\,, \qquad \mathbb{H}^{T}(t',t) = \begin{pmatrix} \Theta(t'-t) & 1 \\ 0 & \Theta(t-t') \end{pmatrix}\,. \tag{5.32}$$

We are now in a position to define a classical closed time path Green function.

Definition 15 (Closed time path Green function)**.** Rather than use the matrix notation for the Green functions, as in Eq. (5.28), we can use the closed time path formalism we developed in Section 3.5 for path integrals in quantum mechanics. The closed time path is the same as shown in Fig. 3.1. In this formulation, the Green function matrix is represented by the location of t and t' on the closed time path contour. The closed time path step function was defined in Eq. (3.37) and explicity given on the contour in Eq. (3.38). On the CTP contour, the complete Green function is given by

$$G(t,t') = \Theta_{\mathcal{C}}(t,t')\, G_{>}(t,t') + \Theta_{\mathcal{C}}(t',t)\, G_{<}(t,t')\,. \tag{5.33}$$

Remark 13. At $t = t'$,

$$G(t,t) = F(t,t)\,. \tag{5.34}$$

Example 25. Let us work out the closed time path Green function for a harmonic oscillator. The Lagrangian is

$$L(q,\dot{q}) = \frac{1}{2}\, m\left[\,\dot{q}^2 - \omega_0^2\, q^2\,\right]. \tag{5.35}$$

Equations of motion are given by

$$\frac{\mathrm{d}^2\langle\, q(t)\,\rangle}{\mathrm{d}t^2} + \omega^2\,\langle\, q(t)\,\rangle = 0\,, \tag{5.36}$$

which have solutions given by

$$q(t) = q\,\cos(\omega t) + (p/m)\,\sin(\omega t)\,. \tag{5.37}$$

So then the correlation function is

$$F(t,t') = \langle\, q(t)\, q(t')\,\rangle = \langle\, q^2\,\rangle\,\cos(\omega t)\,\cos(\omega t') + \langle\, p^2\,\rangle\,\sin(\omega t)\,\sin(\omega t')/m^2 + \langle\, qp\,\rangle\,\sin[\,\omega(t+t')\,]/m\,, \tag{5.38}$$

where

$$\langle\, q^2\,\rangle = \iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\rho(q,p)\,q^2\,, \qquad \langle\, qp\,\rangle = \iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\rho(q,p)\,q\,p\,, \qquad \langle\, p^2\,\rangle = \iint \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\rho(q,p)\,p^2\,. \tag{5.39}$$

The spectral function is given by

$$\begin{aligned}\sigma(t,t') = \langle\,\{\, q(t), q(t')\,\}\,\rangle &= \langle\,\{\, q, q\,\}\,\rangle\,\cos(\omega t)\,\cos(\omega t') + \langle\,\{\, q, p\,\}\,\rangle\,\cos(\omega t)\,\sin(\omega t')/m \\ &\quad + \langle\,\{\, p, q\,\}\,\rangle\,\sin(\omega t)\,\cos(\omega t')/m + \langle\,\{\, p, p\,\}\,\rangle\,\sin(\omega t)\,\sin(\omega t')/m^2 \\ &= -\sin[\,\omega(t-t')\,]/m\,.\end{aligned} \tag{5.40}$$

From these two expressions, we can find all the Green functions. We get

5.1.4 Classical generating functional

In this section, we derive the classical generating functional for closed-time-path Green functions.

5.2 Quantum theory

In quantum mechanics, the density operator $\rho(t)$ is defined in the Schrödinger picture by the outer product of the Schrödinger state vector $|\,\psi(t)\,\rangle$

$$\rho(t) = |\,\psi(t)\,\rangle\langle\,\psi(t)\,| = U(t)\,|\,\psi_0\,\rangle\langle\,\psi_0\,|\,U^{\dagger}(t) = U(t)\,\rho_0\,U^{\dagger}(t)\,, \tag{5.41}$$

where $\rho_0 = |\,\psi_0\,\rangle\langle\,\psi_0\,|$ is the Heisenberg density operator at $t = 0$. $U(t)$ and $U^{\dagger}(t)$ are the time-development operators, given by Eqs. (4.18) and (4.19). The density operator satisfies an equation of motion

$$\frac{\partial\rho(t)}{\partial t} = \frac{\partial|\,\psi(t)\,\rangle}{\partial t}\,\langle\,\psi(t)\,| + |\,\psi(t)\,\rangle\,\frac{\partial\langle\,\psi(t)\,|}{\partial t} = -[\,\rho(t), H\,]/(i\hbar)\,, \tag{5.42}$$

or

$$\frac{\partial \rho(t)}{\partial t} + [\,\rho(t), H\,]/(i\hbar) = 0\,. \tag{5.43}$$

where we have used Schrödinger's equation. Eq. (5.43) is the quantum statement of the classical Liouville theorem, Eq. (5.2).

In a basis $|\,e_i\,\rangle$ of the system, the density matrix $\rho_{ij}(t)$ is given by

$$\rho_{ij}(t) = \langle\, e_i \,|\, \psi(t)\,\rangle\langle\, \psi(t) \,|\, e_j\,\rangle\,. \tag{5.44}$$

The density matrix is normalized and idempotent at all times

$$\mathrm{Tr}[\,\rho(t)\,] = \sum_i \rho_{ii}(t) = \sum_i |\,\langle\, e_i \,|\, \psi(t)\,\rangle\,|^2 = \mathrm{Tr}[\,\rho_0\,] = 1\,, \qquad \rho^2(t) = \rho(t)\,, \tag{5.45}$$

which expresses the conservation of probability.

Let us examine some of the properties of the Heisenberg density matrix at $t = 0$. First of all, $\rho = |\,\psi\,\rangle\langle\,\psi\,|$ is Hermitian, and therefore has an eigenvalue problem which we write as

$$\rho\,|\,\rho_n\,\rangle = \rho_n\,|\,\rho_n\,\rangle\,, \qquad \text{with} \qquad \langle\,\rho_n\,|\,\rho_{n'}\,\rangle = \delta_{n,n'}\,, \qquad \sum_n |\,\rho_n\,\rangle\langle\,\rho_n\,| = 1\,. \tag{5.46}$$

But since the density matrix is idempotent, the eigenvalues must obey the equation: $\rho_n(\rho_n - 1) = 0$, so eigenvalues must be either zero or one: $\rho_n = 0, 1$, for all n. However in addition, the trace of ρ is one, so that the sum of all eigenvalues must also be one: $\sum_n \rho_n = 1$. This means that there can only be one eigenvalue with value one, all the others must be zero. Given a vector $|\,\psi\,\rangle$, we can always construct a density operator $\rho = |\,\psi\,\rangle\langle\,\psi\,|$ which contains all the information in the ray $|\,\psi\,\rangle$, without the arbitrary phase factor associated with the vector $|\,\psi\,\rangle$.

As in the classical case, when we want to find average values of operators of the form,

$$\begin{aligned}\langle\, F(Q(t),P(t))\,\rangle &= \langle\,\psi_0\,|\,F(Q(t),P(t))\,|\,\psi_0\,\rangle = \langle\,\psi(t)\,|\,F(Q,P)\,|\,\psi(t)\,\rangle \\ &= \mathrm{Tr}[\,\rho(t)\,F(Q,P)\,] = \mathrm{Tr}[\,\rho_0\,F(Q(t),P(t))\,]\,,\end{aligned} \tag{5.47}$$

we have our choice of either solving the equations of motion for the operators or the equation of motion (5.43) for the density matrix. They both give the same answer. In many cases in non-relativistic quantum mechanics, the simplest method may be to just solve Schrödinger's equation and then find $\rho(t)$. However if the system has a large number of canonical variables (for example, more than three!), Schrödinger's equation can be very difficult, if not impossible, to solve, and one is forced to look at solutions of the equations of motion in the Heisenberg representation. This is the case for quantum field theory where an infinite and continuous number of canonical variables are needed to describe the physics. So we consider here in this chapter methods that can be used in the Heisenberg representation.

Let us first examine representations of the density matrix. In a coordinate or momentum representation, there are four different density matrices we can define. They are given by the following:

$$\langle\, q\,|\,\rho(t)\,|\,q'\,\rangle\,, \qquad \langle\, q\,|\,\rho(t)\,|\,p\,\rangle\,, \qquad \langle\, p\,|\,\rho(t)\,|\,q\,\rangle\,, \qquad \langle\, p\,|\,\rho(t)\,|\,p'\,\rangle\,. \tag{5.48}$$

But these are all related to each other by Fourier transforms, so if we find one of them we can find them all. Here we will study the density matrix in a coordinate representation given by the first matrix element of the above list, and define

$$\begin{aligned}\rho(q,q',t) &= \langle\, q\,|\,\rho(t)\,|\,q'\,\rangle = \langle\, q\,|\,\psi(t)\,\rangle\,\langle\,\psi(t)\,|\,q'\,\rangle = \langle\, q,t\,|\,\psi_0\,\rangle\,\langle\,\psi_0\,|\,q',t\,\rangle = \langle\, q,t\,|\,\rho_0\,|\,q',t\,\rangle \\ &= \iint \mathrm{d}q''\,\mathrm{d}q'''\,\langle\, q,t\,|\,q'',0\,\rangle\,\rho_0(q'',q''')\,\langle\, q''',0\,|\,q',t\,\rangle\,,\end{aligned} \tag{5.49}$$

where $\rho_0(q'',q''') = \langle\, q \,|\, \rho_0 \,|\, q' \,\rangle = \langle\, q \,|\, \psi_0 \,\rangle\langle\, \psi_0 \,|\, q' \,\rangle$. From the result in (5.49), we see that in order to find the full density matrix in the coordinate representation at time t, we will need to find the propagator,

$$\langle\, q''', 0 \,|\, q', t \,\rangle \langle\, q, t \,|\, q'', 0 \,\rangle \,. \tag{5.50}$$

Here $\langle\, q, t \,|\, q'', 0 \,\rangle$ propagates the system forward in time from a point q'' at $t = 0$ to a point q at time t and then $\langle\, q''', 0 \,|\, q', t \,\rangle$ propagates the system backward in time from a point q' at time t to a point q at time $t = 0$. So *both* propagation forward in time and then backward in time are necessary in order to find $\rho(q, q', t)$.

Normalization of the density matrix in the coordinate representation is given by an integral over the diagonal elements,

$$\int_{-\infty}^{+\infty} \mathrm{d}q\, \rho(q, q, t) = \langle\, \psi(t) \,|\, \psi(t) \,\rangle = 1 \,, \tag{5.51}$$

for all t. The average value of the Heisenberg position operator $Q(t)$ is given by a trace over the density matrix

$$\begin{aligned} \langle\, Q(t) \,\rangle &= \langle\, \psi(t) \,|\, Q \,|\, \psi(t) \,\rangle = \int_{-\infty}^{+\infty} \mathrm{d}q\, \rho(q, q, t)\, q \\ &= \iiint \mathrm{d}q\, \mathrm{d}q'\, \mathrm{d}q''\, \rho_0(q', q'')\, q\, \langle\, q'', 0 \,|\, q, t \,\rangle \langle\, q, t \,|\, q', 0 \,\rangle \,. \end{aligned} \tag{5.52}$$

So in order to calculate this quantity, we will need to find the propagator

$$\langle\, q'', 0 \,|\, q, t \,\rangle \langle\, q, t \,|\, q', 0 \,\rangle \,, \tag{5.53}$$

where we must find the propagator from a point q' at time $t = 0$ to a point q at time t and then from this point back to a point q'' at $t = 0$. We show how to find this propagator in terms of a path integral in Chapter 3. Finding this propagator is the key to obtaining the correlation and Green functions for the system.

References

Chapter 6

Thermal densities

In this chapter, we discuss systems in thermal equilibrium. We study methods to calculate properties of such systems using methods we have developed for quantum mechanics, and employing some of the same ideas.

We start in Section 6.1 by deriving the canonical ensemble for a physical system in quantum mechanics. Then in Section 6.2, we discuss thermodynamic averages of quantum operators. In Section 6.3, we discuss the imaginary time, or Matsubara formalism, and proceed to find Green functions and path integrals for this formalism. In Section 6.6, we discuss the thermovariables method.

6.1 The canonical ensemble

The thermal density matrix ρ for a canonical ensemble is defined to be the normalized operator which minimizes the entropy such that the average energy is constrained to be a fixed number. The definition for the entropy (S) in terms of the thermal density matrix is given by Boltzmann's famous formula,

$$S = -k_{\mathrm{B}} \,\mathrm{Tr}[\,\rho\, \ln[\,\rho\,]\,]\,, \tag{6.1}$$

where k_{B} is Boltzmann's constant. Think of the entropy as measuring the degree of uncertainty of the system. The energy (E) and normalization is given by

$$E = \mathrm{Tr}[\,\rho\, H\,]\,, \qquad 1 = \mathrm{Tr}[\,\rho\,]\,. \tag{6.2}$$

Minimization of S with these constraints gives the canonical density matrix,

$$\rho = \frac{1}{Z}\, e^{-\beta H}\,, \tag{6.3}$$

where Z and β are Lagrange multipliers. Z is fixed in terms of β by the normalization requirement,

$$Z(\beta) = e^{-\beta\Omega(\beta)} = \mathrm{Tr}[\,e^{-\beta H}\,]\,, \tag{6.4}$$

which defines the *grand potential* $\Omega(\beta)$. The entropy is then given by

$$S/k_{\mathrm{B}} = -\mathrm{Tr}[\,\rho\, \ln[\,\rho\,]\,] = \beta\, E + \ln[\,Z(\beta)\,] = \beta\,[\,E - \Omega(\beta)\,]\,. \tag{6.5}$$

The system we are describing may or may not involve something like a gas of particles contained in a fixed volume. If it does, however, we have available the combined first and second laws of thermodynamics, which states that

$$T\,\mathrm{d}S(E,V) = \mathrm{d}E + p\,\mathrm{d}V\,, \tag{6.6}$$

and which defines the temperature (T) and pressure (p). The partition function $Z(\beta, V)$ then depends on V as well as β. From (6.5) and (6.6) we find the partial differential relations,

$$\left[\frac{\partial S(E,V)}{\partial E}\right]_V = \frac{1}{T} = k_{\rm B}\,\beta\,, \qquad \left[\frac{\partial S(E,V)}{\partial V}\right]_E = \frac{p}{T} = -k_{\rm B}\,\beta\left[\frac{\partial \Omega(\beta,V)}{\partial V}\right]_\beta . \tag{6.7}$$

So we find that

$$\beta = 1/(k_{\rm B}T)\,, \qquad \text{and} \qquad p = -\left[\frac{\partial \Omega(\beta,V)}{\partial V}\right]_\beta . \tag{6.8}$$

Even if we cannot define a volume for the system, we can still use the relation $T\,{\rm d}S = {\rm d}E$ to define what we call the "temperature" of the system.

The ensemble average is connected to time-averages by the ergodic hypothesis, which states that

ergodic hypothesis here!

6.2 Ensemble averages

In 1932, Felix Block [1] proposed that the ensemble average of any quantum mechanical operator A in the Schrödinger representation is given by

$$\langle\, A\,\rangle_\beta = {\rm Tr}[\,\rho(\beta)\,A\,]\,, \qquad \text{where} \qquad \rho(\beta) = \frac{1}{Z(\beta)}\,e^{-\beta H}\,. \tag{6.9}$$

This prescription is the same that we used in Chapter 5 for average values, except that here $\rho(\beta)$ is an ensemble density matrix rather than the density matrix of the quantum state of the system. Let us be clear that it is impossible to write the canonical ensemble density operator as the outer product of some vector in a Hilbert space, so that $\rho(\beta)$ is not a density operator describing a state of the system! Rather we should think of it as describing an average state of the system which minimizes the entropy, or degree of uncertainty. We state this in the form of a theorem in the following

Theorem 13. *There is no vector* $|\,\psi(\beta)\,\rangle$ *such that*

$$|\,\psi(\beta)\,\rangle\langle\,\psi(\beta)\,| = \rho(\beta) = \frac{1}{Z(\beta)}\,e^{-\beta H} \tag{6.10}$$

except for a possible trivial case.

Proof. The proof is easy and left as an exercise. □

6.3 Imaginary time formalism

The factor $\exp[\,-\beta H\,]$ in the density matrix for the canonical ensemble is very suggestive of the time development operator $U(t) = \exp[\,-iHt/\hbar\,]$ in quantum mechanics for negative complex time. In fact, if we put

$$t/\hbar \mapsto -i\tau\,, \tag{6.11}$$

we find

$$T(\tau) = U(-i\hbar\,\tau) = e^{-\tau H}\,. \tag{6.12}$$

Here $T(\tau)$ is an invertable *Hermitian* operator, not a unitary transformation for τ real and positive. So lengths and angles are *not* preserved by this transformation. However we can still use (6.12) to define "thermal Schrödinger" and "thermal Heisenberg" pictures. Let us put

$$|\,\psi(\tau)\,\rangle = T(\tau)\,|\,\psi\,\rangle\,, \qquad \text{and} \qquad Q(\tau) = T^{-1}(\tau)\,Q\,T(\tau)\,, \qquad P(\tau) = T^{-1}(\tau)\,P\,T(\tau)\,, \tag{6.13}$$

for any vector $|\,\psi\,\rangle$ and operators Q and P in the Schrödinger picture. For any function of Q and P, we have

$$F(\,Q(\tau),P(\tau)\,) = T^{-1}(\tau)\,F(Q,P)\,T(\tau)\,. \tag{6.14}$$

In particular, if $F(Q,P) = H(Q,P)$, we have

$$H(\,Q(\tau),P(\tau)\,)T = T^{-1}(\tau)\,H(Q,P)\,T(\tau) = H(Q,P)\,, \tag{6.15}$$

since $[\,T(\tau),H\,] = 0$. The thermal vector $|\,\psi(\tau)\,\rangle$ satisfies a "thermal Schrödinger" equation,

$$\frac{\mathrm{d}|\,\psi(\tau)\,\rangle}{\mathrm{d}\tau} = -H\,|\,\psi(\tau)\,\rangle\,. \tag{6.16}$$

Here we constrain the imaginary time variable τ to be in the range $0 \le \tau \le \beta$, so that a formal solution of the thermal Schrödinger equation (6.16) is given by

$$|\,\psi(\beta)\,\rangle = \exp\left\{-\int_0^\beta H\,\mathrm{d}\tau\right\}|\,\psi\,\rangle = T(\beta)\,|\,\psi\,\rangle\,, \tag{6.17}$$

for τ-independent thermal Hamiltonians. Eq. (6.17) maps all state vectors $|\,\psi\,\rangle$ in Hilbert space to thermal vectors $|\,\psi(\beta)\,\rangle$, along a path governed by the thermal Schrödinger equation. $Q(\tau)$ and $P(\tau)$ satisfy "thermal Heisenberg" equations of motion,

$$\frac{\mathrm{d}Q(\tau)}{\mathrm{d}\tau} = -[\,Q(\tau),H\,]\,, \qquad \frac{\mathrm{d}P(\tau)}{\mathrm{d}\tau} = -[\,P(\tau),H\,]\,, \tag{6.18}$$

and obey the equal τ commutation relations,

$$[\,Q(\tau),P(\tau)\,] = T^{-1}(\tau)\,[\,Q,P\,]\,T(\tau) = i\hbar\,. \tag{6.19}$$

For Hamiltonians of the form $H = P^2/(2m) + V(Q)$, the thermal Heisenberg equations of motion are

$$\frac{\mathrm{d}Q(\tau)}{\mathrm{d}\tau} = -[\,Q(\tau),P^2(\tau)/(2m)\,] = -i\hbar\,P(\tau)/m\,, \tag{6.20a}$$

$$\frac{\mathrm{d}P(\tau)}{\mathrm{d}\tau} = -[\,P(\tau),V(\,Q(\tau)\,)\,]\,. \tag{6.20b}$$

So from (6.20a), we have

$$P(\tau) = \frac{i}{\hbar}\,m\,Q'(\tau)\,. \tag{6.21}$$

Here we use a prime to indicate differentiation with respect to τ.

We now come to an important result. According to the Block prescription (6.9) for finding thermal averages, the thermal averages of the operators $Q(\tau)$ and $P(\tau)$ are periodic with period β. We state this in a general form in the following theorem:

Theorem 14. *The thermal average of any function $F(\,Q(\tau),P(\tau)\,)$ is periodic in τ with period β,*

$$\langle\,F(\,Q(\tau+\beta),P(\tau+\beta)\,)\,\rangle_\beta = \langle\,F(\,Q(\tau),P(\tau)\,)\,\rangle_\beta\,. \tag{6.22}$$

Proof. We find

$$\begin{aligned}
\langle\,F(\,Q(\tau+\beta),P(\tau+\beta)\,)\,\rangle_\beta &= \frac{1}{Z(\beta)}\,\mathrm{Tr}[\,e^{-\beta H}\,F(\,Q(\tau+\beta),P(\tau+\beta)\,)\,] \\
&= \frac{1}{Z(\beta)}\,\mathrm{Tr}[\,e^{-\beta H}\,e^{(\tau+\beta)H}\,F(Q,P)\,e^{-(\tau+\beta)H}\,] \\
&= \frac{1}{Z(\beta)}\,\mathrm{Tr}[\,e^{\tau H}\,F(Q,P)\,e^{-(\tau+\beta)H}\,] = \frac{1}{Z(\beta)}\,\mathrm{Tr}[\,e^{-\beta H}\,e^{\tau H}\,F(Q,P)\,e^{-\tau H}\,] \\
&= \frac{1}{Z(\beta)}\,\mathrm{Tr}[\,e^{-\beta H}\,F(\,Q(\tau),P(\tau)\,)\,] = \langle\,F(\,Q(\tau),P(\tau)\,)\,\rangle_\beta\,,
\end{aligned} \tag{6.23}$$

which is what we were trying to prove. □

In particular, we have

$$\langle\, Q(\tau+\beta)\,\rangle_\beta = \langle\, Q(\tau)\,\rangle_\beta\,, \qquad \langle\, P(\tau+\beta)\,\rangle_\beta = \langle\, P(\tau)\,\rangle_\beta\,. \tag{6.24}$$

This means that we can expand $\langle\, Q(\tau)\,\rangle_\beta$ and $\langle\, P(\tau)\,\rangle_\beta$ in a fourier series with period β,

$$\langle\, Q(\tau)\,\rangle_\beta = \frac{1}{\beta}\sum_{n=-\infty}^{+\infty} Q_n\, e^{-i2\,\omega_n\tau}\,, \qquad \text{and} \qquad \langle\, P(\tau)\,\rangle_\beta = \frac{1}{\beta}\sum_{n=-\infty}^{+\infty} P_n\, e^{-i2\,\omega_n\tau}\,. \tag{6.25}$$

where the frequencies ω_n are given by

$$2\,\omega_n = 2\pi\, n/\beta\,. \tag{6.26}$$

(The factor of two in these definitions are explained below.) For Hamiltonians of the form, $H = P^2/(2m) + V(Q)$, we have

$$P_n = \frac{m\,\omega_n}{\hbar}\,Q_n\,. \tag{6.27}$$

Theorem 14 states that the *thermal averages* of thermal Heisenberg operators are periodic. However, we cannot conclude that the thermal operators themselves are periodic. In fact the operators depend on τ, not β.

We define τ-ordered products exactly like the time-ordered ones. We put

$$\mathcal{T}_\tau\{\, Q(\tau), Q(\tau')\,\} = Q(\tau)\,Q(\tau')\,\Theta(\tau-\tau') + Q(\tau')\,Q(\tau)\,\Theta(\tau'-\tau)\,, \tag{6.28}$$

and define a thermal two-point Green function by

$$G(\tau,\tau') = i\,\langle\, \mathcal{T}_\tau\{\, Q(\tau), Q(\tau')\,\}\,\rangle_\beta/\hbar = G_>(\tau,\tau')\,\Theta(\tau-\tau') + G_<(\tau,\tau')\,\Theta(\tau'-\tau)\,, \tag{6.29}$$

where

$$G_>(\tau,\tau') = i\,\langle\, Q(\tau)\,Q(\tau')\,\rangle_\beta/\hbar\,, \tag{6.30a}$$
$$G_<(\tau,\tau') = i\,\langle\, Q(\tau')\,Q(\tau)\,\rangle_\beta/\hbar\,. \tag{6.30b}$$

Let us first note that $G_>(\tau,\tau')$ and $G_<(\tau,\tau')$ are functions of $\tau-\tau'$, since, for example, we can write

$$\begin{aligned}\langle\, Q(\tau)\,Q(\tau')\,\rangle_\beta &= \frac{1}{Z(\beta)}\,\mathrm{Tr}[\, e^{-\beta H}\,Q(\tau)\,Q(\tau')\,] \\ &= \frac{1}{Z(\beta)}\,\mathrm{Tr}[\, e^{-\beta H}\,e^{\tau H} Q\, e^{-(\tau-\tau')H}\,Q\, e^{-\tau' H}\,] \\ &= \frac{1}{Z(\beta)}\,\mathrm{Tr}[\, e^{-\beta H}\,e^{(\tau-\tau')H} Q\, e^{-(\tau-\tau')H}\,Q\,]\,,\end{aligned} \tag{6.31}$$

which is a function of $\tau-\tau'$. So let us put $\tau'=0$, and write

$$G(\tau) = G_>(\tau)\,\Theta(\tau) + G_<(\tau)\,\Theta(-\tau)\,, \tag{6.32}$$

where now

$$G_>(\tau) = i\,\langle\, Q(\tau)\,Q(0)\,\rangle_\beta/\hbar\,, \tag{6.33a}$$
$$G_<(\tau) = i\,\langle\, Q(0)\,Q(\tau)\,\rangle_\beta/\hbar\,. \tag{6.33b}$$

Note that $G_<(\tau) = G_>(-\tau)$ so that $G(-\tau) = G(\tau)$, and is an even function of τ. The next theorem, due to Kubo[2], and Martin and Schwinger[3], is similar to Theorem 14 above and relates $G_>(\tau+\beta)$ to $G_<(\tau)$.

Theorem 15 (KMS theorem)**.** *The theorem states that*

$$G_>(\tau+\beta) = G_<(\tau)\,. \tag{6.34}$$

Proof. We find that

$$\begin{aligned}\langle\, Q(\tau+\beta)\, Q(0)\,\rangle_\beta &= \frac{1}{Z(\beta)}\,\mathrm{Tr}[\, e^{-\beta H}\, Q(\tau+\beta)\, Q(0)\,] \\ &= \frac{1}{Z(\beta)}\,\mathrm{Tr}[\, e^{\tau H}\, Q\, e^{-(\tau+\beta)H} Q\,] \\ &= \frac{1}{Z(\beta)}\,\mathrm{Tr}[\, e^{-\beta H}\, Q\, e^{\tau H}\, Q\, e^{-\tau H}\,] = \langle\, Q(0)\, Q(\tau)\,\rangle_\beta\,.\end{aligned} \tag{6.35}$$

The result now follows from the definitions (6.33). □

The KMS theorem is only one of a number of similar theorems. Using the KMS theorem, we find that

$$G(\beta) = G_>(\beta) = G_>(0) = G_<(0) = G_<(-\beta) = G(-\beta)\,. \tag{6.36}$$

In other words the argument of $G(\tau)$ is in the range $-\beta \le \tau \le +\beta$, and is even, with boundary conditions such that $G(-\beta) = G(+\beta)$. This means that we can expand $G(\tau)$ in a fourier series given by

$$G(\tau) = \frac{1}{\beta}\sum_{n=-\infty}^{+\infty} G_n\, e^{-i\omega_n\tau}\,, \tag{6.37a}$$

$$G_n = \frac{1}{2}\int_{-\beta}^{+\beta} \mathrm{d}\tau\, G(\tau)\, e^{+i\omega_n\tau} = \int_0^{+\beta} \mathrm{d}\tau\, G(\tau)\, e^{+i\omega_n\tau}\,, \tag{6.37b}$$

where the Matsubara frequencies ω_n are given by[4]

$$\omega_n = \pi n/\beta\,. \tag{6.38}$$

Notice that these frequencies are one-half the frequencies found for the expansions of $\langle\, Q(\tau)\,\rangle_\beta$ and $\langle\, P(\tau)\,\rangle_\beta$ in Eq. (6.26).

6.4 Thermal Green functions

General imaginary time Green functions are defined in a way analogous to the real time case. We put

$$\tau\{\, Q(\tau_1), Q(\tau_2), \ldots, Q(\tau_n)\,\}_\beta = \langle\, \mathcal{T}_\tau\{\, Q(\tau_1), Q(\tau_2), \ldots, Q(\tau_n)\,\}\,\rangle_\beta\,. \tag{6.39}$$

6.5 Path integral representation

From our discussion of path integrals in Chapter 3, we found that the propagator $\langle\, q,t\,|\,q,0\,\rangle$ could be written as a path integral given by

$$\langle\, q\,|\,U(t)\,|\,q'\,\rangle = \mathcal{N}\int_{q(0)=q'}^{q(t)=q} \mathrm{D}q\, \exp\left\{\frac{i}{\hbar}\int_0^t \mathrm{d}t'\left[\frac{1}{2}m\,\dot{q}^2 + V(q)\right]\right\} \tag{6.40}$$

Translating this expression to imaginary time according to Eq. (6.11), $t/\hbar \mapsto -i\beta$, we find

$$\langle\, q\,|\,e^{-\beta H}\,|\,q'\,\rangle = \mathcal{N}\int_{q(0)=q'}^{q(\beta)=q} \mathrm{D}q\, \exp[\,-S_\mathrm{E}[\,q\,]\,]\,, \tag{6.41}$$

where $S_\mathrm{E}[\,q\,]$ is the Euclidean action

$$S_\mathrm{E}[\,q\,] = \int_0^\beta \mathrm{d}\beta'\left[\frac{1}{2}m\,q'^2 - V(q)\right]. \tag{6.42}$$

Here we have mapped

$$q(t) \mapsto q(\beta)\,, \qquad \text{and} \qquad \dot{q}(t) \mapsto i\,q' = i\frac{\mathrm{d}q}{\mathrm{d}\beta}\,. \tag{6.43}$$

where $L_{\mathrm{E}}[\phi]$ is the Euclidean Lagrangian. Eq. (**??**) then becomes:

$$\langle\,\phi(x)\,|\,e^{-\beta H}\,|\,\phi(x')\,\rangle = \bar{N}\int_{\phi(x',0)}^{\phi(x,\beta)}\mathcal{D}\phi\,e^{-S_{\mathrm{E}}[\phi]} \tag{6.44}$$

6.6 Thermovariable methods

Now let us consider the possibility of a density matrix at $t = 0$ of the canonical form:

$$\rho(\beta, Q, P) = e^{-\beta H(Q,P)}/Z(\beta)\,, \tag{6.45}$$

where $H(Q, P)$ is the Hamiltonian for the particle. $Z(\beta)$ is chosen to normalize the trace of ρ to one:

$$\mathrm{Tr}[\,\rho(\beta, Q, P)\,] = 1\,, \qquad \Rightarrow \qquad Z(\beta) = \mathrm{Tr}[\,e^{-\beta H(Q,P)}\,]\,, \tag{6.46}$$

We see immediately that $\rho(\beta, Q, P)$ is Hermitian and has unit trace, but it is not idempotent: $\rho^2(\beta, Q, P) \neq \rho(\beta, Q, P)$, so it is impossible to find a vector $|\,\psi\,\rangle$ such that $\rho(\beta, Q, P) = |\,\psi\,\rangle\langle\,\psi\,|$. We can see this in another way. Since the Hamiltonian obeys an eigenvalue problem,

$$H(Q, P)\,|\,E_n\,\rangle = E_n\,|\,E_n\,\rangle\,, \tag{6.47}$$

we see that

$$\langle\,E_n\,|\,\rho\,|\,E_{n'}\,\rangle = e^{-\beta E_n}\,\delta_{n,n'}/Z(\beta) = \langle\,E_n\,|\,\psi\,\rangle\,\langle\,\psi\,|\,E_{n'}\,\rangle = \psi_{E_n}\,\psi^*_{E_{n'}}\,. \tag{6.48}$$

However ψ_{E_n} is just a complex number, so there is no way to satisfy Eq. (6.48), since in general $H(Q, P)$ has more than one eigenvalue. So it appears impossible to choose ρ to be a statistical state. However, we notice that Eq. (6.48) looks like an orthogonal requirement for state vectors, but ψ_{E_n} are not state vectors. The way out of this is to double the Hilbert space and introduce a second eigenvector. This method is called "Thermofield Dynamics," and was invented by Kubo and by Martin and Schwinger in the late '50's. So we put our Hilbert space as consisting of the direct sum: $\mathcal{H}(Q, P) = H(Q, P) \oplus H(Q, P)$ and the vectors as direct products: $|\,n, m\,\rangle = |\,E_n\,\rangle \otimes |\,E_m\,\rangle$. So any operator, including the density matrix, is also a direct product. The first system does not act on the second, so that, for example:

$$\begin{aligned}\langle\,n, m\,|\,A(Q, P)\otimes 1\,|\,n', m'\,\rangle &= \langle\,n\,|\,A(Q, P)\,|\,n'\,\rangle\,\delta_{m,m'}\,,\\ \langle\,n, m\,|\,1\otimes A(Q, P)\,|\,n', m'\,\rangle &= \langle\,m\,|\,A(Q, P)\,|\,m'\,\rangle\,\delta_{n,n'}\,.\end{aligned} \tag{6.49}$$

This kind of behavior is just what we need to satisfy Eq. (6.48). We can define a state $|\,\psi(\beta)\,\rangle$ as follows:

$$|\,\psi(\beta)\,\rangle = \sum_n \psi_{E_n}\,|\,n, n\,\rangle\,. \tag{6.50}$$

References

[1] F. Block, Z. Physik **74**, 295 (1932).

[2] R. Kubo, "Statistical mechanical theory of irreversible processes. 1. General theory and simple applications in magnetic and conduction problems," Phys. Soc. Japan **12**, 570 (1957).

[3] P. Martin and J. Schwinger, "Classical perturbation theory," Phys. Rev. **115**, 1342 (1959).

[4] T. Matsubara, "Statistical mechanics theory," Prog. Theo. Physics **14**, 351 (1955).

Chapter 7

Green functions

Here we define quantum mechanical Green functions.

References

Chapter 8

Identical particles

In this chapter, we discuss the quantum mechanics of identical particles. By identical, we mean that the Hamiltonian describing them is invariant under interchange of any two particles. That is, there is no physical property by which we can distinguish them. If we let $\mathbf{r}_i$, $i = 1, 2, \ldots, N$ be the coordinates of a particle, then we require the probability density to be the same under interchange of any two particles. That is:

$$|\,\psi(\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_N, t)\,|^2 = |\,\psi(\mathbf{r}_1, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_N, t)\,|^2\,, \tag{8.1}$$

for all time t. There are only two known solutions for the wave functions, namely:

$$\psi^{(\pm)}(\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_N, t) = \pm\psi^{(\pm)}(\mathbf{r}_1, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_N, t)\,. \tag{8.2}$$

Even wave functions describe what we call particles with Bose statistics, and odd wave functions describe what we call Fermi statistics. There is a connection between the spin of the particles and the type of statistics for that particle, which involves special relativity and is beyond the scope of this book. For non-relativistic particles, theoretically either statistics could apply, but experimentally we observe that particles with integer spin, $S = 0, 2, \ldots$ obey Bose statistics and particles with half-integer spin, $S = 1/2, 3/2, \ldots$ obey Fermi statistics.

8.1 Coordinate representation

We assume that we can describe particles by N independent Cartesian coordinates, $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N$, and canonical momenta, $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N$. In quantum mechanics, these quantities become operators $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_N$ and $\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_N$, with the following commutation properties:

$$[\,R_{i,a}, P_{j,b}\,] = i\hbar\,\delta_{i,j}\delta_{a,b}\,, \qquad [\,R_{i,a}, R_{j,b}\,] = [\,P_{i,a}, P_{j,b}\,] = 0\,. \tag{8.3}$$

Here the middle alphabet Roman letters $i, j, \ldots$ refer to the particle and the beginning alphabet Roman letters $a, b, \ldots$ refer to the Cartesian x, y, z coordinates. Eigenvalue equations for $\mathbf{R}_i$ and $\mathbf{P}_i$ are:

$$\mathbf{R}_i\,|\,\mathbf{r}_i\,\rangle = \mathbf{r}_i\,|\,\mathbf{r}_i\,\rangle\,, \qquad \text{and} \qquad \mathbf{P}_i\,|\,\mathbf{p}_i\,\rangle = \mathbf{p}_i\,|\,\mathbf{p}_i\,\rangle\,. \tag{8.4}$$

Eigenvectors for all the coordinates are then constructed by a direct product:

$$|\,\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N\,\rangle = |\,\mathbf{r}_1\,\rangle \otimes |\,\mathbf{r}_2\,\rangle \otimes \cdots \otimes |\,\mathbf{r}_N\,\rangle\,, \tag{8.5}$$

with a similar relation for the momentum eigenvector. Fully symmetric (Bose) and antisymmetric (Fermi) eigenvectors are constructed by similar direct products. For example, for two particles, we construct symmetric and antisymmetric direct products of the base vectors:

$$|\,\mathbf{r}_1, \mathbf{r}_2\,\rangle^{(\pm)} = \frac{1}{\sqrt{2}}\left\{\,|\,\mathbf{r}_1\,\rangle \otimes |\,\mathbf{r}_2\,\rangle \pm |\,\mathbf{r}_2\,\rangle \otimes |\,\mathbf{r}_1\,\rangle\,\right\}. \tag{8.6}$$

These symmetric or antisymmetric basis states are normalized such that:

$$ {}^{(\pm)}\langle\, \mathbf{r}_1, \mathbf{r}_2 \,|\, \mathbf{r}'_1, \mathbf{r}'_2 \,\rangle^{(\pm)} = \delta(\mathbf{r}_1 - \mathbf{r}'_1)\,\delta(\mathbf{r}_2 - \mathbf{r}'_2) \pm \delta(\mathbf{r}_1 - \mathbf{r}'_2)\,\delta(\mathbf{r}_2 - \mathbf{r}'_1)\,. \tag{8.7} $$

An operator $T(\mathbf{R}_1, \mathbf{R}_2) = T(\mathbf{R}_2, \mathbf{R}_1)$ which is invariant with respect to interchange of the particles has the value $T(\mathbf{r}_1, \mathbf{r}_2)$ when operating on a fully symmetric or antisymmetric base vectors. That is,

$$ \begin{aligned} T(\mathbf{R}_1, \mathbf{R}_2)\,|\, \mathbf{r}_1, \mathbf{r}_2 \,\rangle^{(\pm)} &= \frac{1}{\sqrt{2}} \left\{\, T(\mathbf{R}_1, \mathbf{R}_2)\,|\, \mathbf{r}_1 \,\rangle \otimes |\, \mathbf{r}_2 \,\rangle \pm T(\mathbf{R}_1, \mathbf{R}_2)\,|\, \mathbf{r}_2 \,\rangle \otimes |\, \mathbf{r}_1 \,\rangle \,\right\} \\ &= \frac{1}{\sqrt{2}} \left\{\, T(\mathbf{r}_1, \mathbf{r}_2)\,|\, \mathbf{r}_1 \,\rangle \otimes |\, \mathbf{r}_2 \,\rangle \pm T(\mathbf{r}_2, \mathbf{r}_1)\,|\, \mathbf{r}_2 \,\rangle \otimes |\, \mathbf{r}_1 \,\rangle \,\right\} \\ &= T(\mathbf{r}_1, \mathbf{r}_2)\,|\, \mathbf{r}_1, \mathbf{r}_2 \,\rangle^{(\pm)}\,. \end{aligned} \tag{8.8} $$

In general, the fully symmetric or antisymmetric direct product can be constructed from a perminate or determinate defined as follows:

$$ |\, \mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N \,\rangle^{(\pm)} = \frac{1}{\sqrt{N!}} \sum_{\mathcal{P}}^{N} (\pm)^{\mathcal{P}}\, \mathcal{P} \left\{\, |\, \mathbf{r}_{i_1} \,\rangle \otimes |\, \mathbf{r}_{i_2} \,\rangle \otimes \cdots \otimes |\, \mathbf{r}_{i_N} \,\rangle \,\right\}, \tag{8.9} $$

where the sum is over all permutations $\mathcal{P}$ of the set $\{i_1, i_2, \cdots, i_N\}$ from the standard set $\{1, 2, \ldots, N\}$ of indices, with a sign assigned to even or odd permutations for the case of Fermi statistics. These symmetric or antisymmetric vectors obey the normalization:

$$ {}^{(\pm)}\langle\, \mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N \,|\, \mathbf{r}'_1, \mathbf{r}'_2, \ldots, \mathbf{r}'_N \,\rangle^{(\pm)} = \sum_{\mathcal{P}}^{N} (\pm)^{\mathcal{P}}\, \delta(\mathbf{r}_1 - \mathbf{r}'_1)\,\delta(\mathbf{r}_2 - \mathbf{r}'_2) \cdots \delta(\mathbf{r}_N - \mathbf{r}'_N)\,, \tag{8.10} $$

where the sum is over all permutations of the primed (or unprimed) indices.

The Hamiltonian for N identical particles of mass m interacting with two-particle forces that depend on the distance between them is given by:

$$ H = \sum_{i=1}^{N} \frac{\mathbf{P}_i^2}{2m} + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} V(\mathbf{R}_i - \mathbf{R}_j)\,. \tag{8.11} $$

Each term in this Hamiltonian is invariant under exchange of any two particles. Schrödinger's equation is given by:

$$ H\,|\, \psi(t) \,\rangle = i\hbar\, \frac{\partial}{\partial t}\,|\, \psi(t) \,\rangle\,, \qquad \text{or} \qquad |\, \psi(t) \,\rangle = e^{-iHt/\hbar}\,|\, \psi \,\rangle\,. \tag{8.12} $$

So the multiple particle state vector in the coordinate representation in the Schrödinger or Heisenberg picture can be written as:

$$ \psi^{(\pm)}(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, t) = {}^{(\pm)}\langle\, \mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N \,|\, \psi(t) \,\rangle = {}^{(\pm)}\langle\, \mathbf{r}_1, t; \mathbf{r}_2, t; \ldots\,; \mathbf{r}_N, t \,|\, \psi \,\rangle\,. \tag{8.13} $$

Notice that there is only one time variable in the Heisenberg picture which describe the positions of all the particles at the same time.

8.2 Occupation number representation

In the last section, we used a coordinate basis to describe the particles, however we are free to use any complete basis to describe particles. Let us suppose that $|\, \alpha \,\rangle$ is such a basis and obeys:

$$ \langle\, \alpha \,|\, \beta \,\rangle = \delta(\alpha - \beta)\,, \qquad \text{and} \qquad \sum_{\alpha} |\, \alpha \,\rangle\langle\, \alpha \,| = 1\,, \tag{8.14} $$

and let us define $\phi_\alpha(\mathbf{r}) = \langle\, \mathbf{r}\,|\,\alpha\,\rangle$ as the overlap between the coordinate representation and the $|\,\alpha\,\rangle$ representation. Then we can write:

$$|\,\mathbf{r}\,\rangle = \sum_\alpha |\,\alpha\,\rangle\,\phi^*_\alpha(\mathbf{r})\,, \tag{8.15}$$

for each particle. Now let us invent an occupation number vector $|\,n_\alpha\,\rangle$ for each of the vectors $|\,\alpha\,\rangle$ which is an eigenstate of number operators $A^{(\pm)\,\dagger}_\alpha A^{(\pm)}_\alpha$ such that:

$$A^{(\pm)\,\dagger}_\alpha A^{(\pm)}_\alpha\,|\,n_\alpha\,\rangle = n_\alpha\,|\,n_\alpha\,\rangle\,, \tag{8.16}$$

with $A^{(+)}_\alpha$ obeying the harmonic oscillator commutator algebra and $A^{(-)}_\alpha$ obeying the Fermi oscillator anti-commutator algebra:

$$[\,A^{(+)}_\alpha, A^{(+)\,\dagger}_\beta\,] = \delta_{\alpha,\beta}\,, \qquad [\,A^{(+)}_\alpha, A^{(+)}_\beta\,] = [\,A^{(+)\,\dagger}_\alpha, A^{(+)\,\dagger}_\beta\,] = 0\,, \tag{8.17}$$

$$\{\,A^{(-)}_\alpha, A^{(-)\,\dagger}_\beta\,\} = \delta_{\alpha,\beta}\,, \qquad \{\,A^{(-)}_\alpha, A^{(-)}_\beta\,\} = \{\,A^{(-)\,\dagger}_\alpha, A^{(-)\,\dagger}_\beta\,\} = 0\,, \tag{8.18}$$

so that $n_\alpha = 0, 1, 2, \ldots$ for Bose statistics and $n_\alpha = 0, 1$ for Fermi statistics. We now put:

$$|\,\alpha\,\rangle = |\,1_\alpha\,\rangle = A^{(\pm)\,\dagger}_\alpha\,|\,0\,\rangle\,. \tag{8.19}$$

Then from Eq. (8.15), we find:

$$|\,\mathbf{r}\,\rangle = \sum_\alpha |\,\alpha\,\rangle\,\phi^*_\alpha(\mathbf{r}) = \sum_\alpha \phi^*_\alpha(\mathbf{r})\,A^{(\pm)\,\dagger}_\alpha\,|\,0\,\rangle \equiv \Phi^{(\pm)\,\dagger}(\mathbf{r})\,|\,0\,\rangle\,, \tag{8.20}$$

where we have defined the field operator $\Phi(\mathbf{r})$ by:

$$\Phi^{(\pm)}(\mathbf{r}) = \sum_\alpha A^{(\pm)}_\alpha\,\phi_\alpha(\mathbf{r})\,, \qquad \text{and} \qquad \Phi^{(\pm)\,\dagger}(\mathbf{r}) = \sum_\alpha A^{(\pm)\,\dagger}_\alpha\,\phi^*_\alpha(\mathbf{r})\,, \tag{8.21}$$

which operates in the occupation number space. So from Eq. (8.9), we find:

$$\begin{aligned} |\,\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N\,\rangle^{(\pm)} &= \frac{1}{\sqrt{N!}} \sum_{\mathcal{P}}^{N} (\pm)^{\mathcal{P}}\,\mathcal{P}\big\{\,|\,\mathbf{r}_{i_1}\,\rangle \otimes |\,\mathbf{r}_{i_2}\,\rangle \otimes \cdots \otimes |\,\mathbf{r}_{i_N}\,\rangle\,\big\}\,, \\ &= \frac{1}{\sqrt{N!}}\,\Phi^{(\pm)\,\dagger}(\mathbf{r}_1)\,\Phi^{(\pm)\,\dagger}(\mathbf{r}_2)\,\cdots\,\Phi^{(\pm)\,\dagger}(\mathbf{r}_N)\,|\,0\,\rangle\,. \end{aligned} \tag{8.22}$$

This last expression includes all permutations of the set of coordinates.

8.3 Particle fields

$$\begin{aligned} \psi^{(\pm)}(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, t) &= \frac{1}{\sqrt{N!}}\,\langle\,0\,|\,\Phi^{(\pm)}(\mathbf{r}_1)\,\Phi^{(\pm)}(\mathbf{r}_2)\,\cdots\,\Phi^{(\pm)}(\mathbf{r}_N)\,|\,\psi(t)\,\rangle\,, \\ &= \frac{1}{\sqrt{N!}}\,\langle\,0\,|\,\Phi^{(\pm)}(\mathbf{r}_1, t)\,\Phi^{(\pm)}(\mathbf{r}_2, t)\,\cdots\,\Phi^{(\pm)}(\mathbf{r}_N, t)\,|\,\psi\,\rangle\,, \end{aligned} \tag{8.23}$$

8.3.1 Hamiltonian

The second quantized Hamiltonian for a system of identical particles is of the form:

$$H = -\int \mathrm{d}^3 r\,\Phi^\dagger(\mathbf{r})\left(\frac{\hbar^2\,\nabla^2}{2m}\,\Phi(\mathbf{r})\right) + \frac{1}{2}\iint \mathrm{d}^3 r\,\mathrm{d}^3 r'\,\Phi^\dagger(\mathbf{r})\,\Phi^\dagger(\mathbf{r}')\,V(\mathbf{r}-\mathbf{r}')\,\Phi(\mathbf{r}')\,\Phi(\mathbf{r}')\,. \tag{8.24}$$

References

Chapter 9

Space-time symmetry transformations

In the last chapter, we set up a vector space which we will use to describe the state of a system of physical particles. In this chapter, we investigate the requirements of space-time symmetries that must be satisfied by a theory of matter. For particle velocities small compared to the velocity of light, the classical laws of nature, governing the dynamics and interactions of these particles, are invariant under the Galilean group of space-time transformations. It is natural to assume that quantum dynamics, describing the motion of non-relativistic particles, also should be invariant under Galilean transformations.

Galilean transformation are those that relate events in two coordinate systems which are spatially rotated, translated, and time-displaced with respect to each other. The invariance of physical laws under Galilean transformations insure that no physical device can be constructed which can distinguish the difference between these two coordinate systems. So we need to assure that this symmetry is built into a non-relativistic quantum theory of particles: we must be unable, by any measurement, to distinguish between these coordinate systems. More generally, a symmetry transformation is a change in state that does not change the results of possible experiments. We formulate this statment in the form of a **relativity principle**:

Definition 16 (Relativity principle)**.** If $|\,\psi(\Sigma)\,\rangle$ represents the state of the system which refers to coordinate system Σ, and if $a(\Sigma)$ is the value of a possible observable operator $A(\Sigma)$ with eigenvector $|\,a(\Sigma)\,\rangle$, also referring to system Σ, then the *probability* $\mathcal{P}_a$ of observing this measurement in coordinate system Σ must be the same as the probability $\mathcal{P}'_a$ of observing this measurement in system Σ', where Σ' is related to Σ by a Galilean transformation. That is, the relativity principle requires that:

$$\mathcal{P}'_a = |\langle\, a(\Sigma')\,|\,\psi(\Sigma')\,\rangle|^2 = \mathcal{P}_a = |\langle\, a(\Sigma)\,|\,\psi(\Sigma)\,\rangle|^2\,. \tag{9.1}$$

In quantum theory, transformations between coordinate systems are written in as operators acting on vectors in $\mathcal{V}$. So let

$$|\,\psi(\Sigma')\,\rangle = U(G)\,|\,\psi(\Sigma)\,\rangle\,, \qquad \text{and} \qquad |\,a(\Sigma')\,\rangle = U(G)\,|\,a(\Sigma)\,\rangle\,, \tag{9.2}$$

where $U(G)$ is the operator representing a Galilean transformation between Σ' and Σ. Then a theorem by Wigner[1] states that:

Theorem 16 (Wigner)**.** *Transformations between two rays in Hilbert space which preserve the same probabilities for experiments are either unitary and linear or anti-unitary and anti-linear.*

Proof. We can easily see that if $U(G)$ is either unitary or anti-unitary, the statement is true. The reverse proof that this is the *only* solution is lengthy, and we refer to Weinberg [**?**][see Weinberg, Appendix A, p. 91] for a careful proof. □

The group of rotations and space and time translations which can be evolved from unity are linear unitary transformations. Space and time reversals are examples of anti-linear and anti-unitary transformations. We will deal with the anti-linear symmetries later on in this chapter.

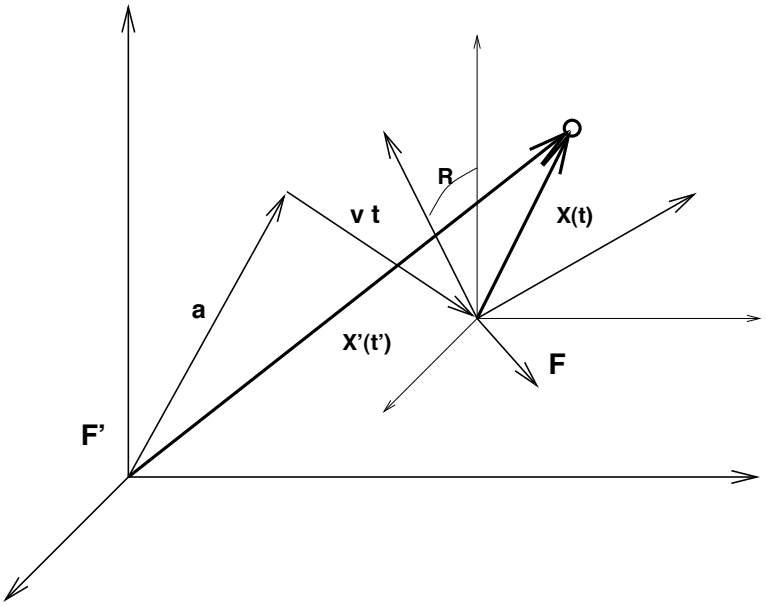

Figure 9.1: The Galilean transformation for Eq. (9.1).

We start this chapter by learning how to describe Galilean transformations in quantum mechanics, and how to classify vectors in Hilbert space according to the way they transform under Galilean transformations. In the process, we will obtain a description of matter, based on the irreducible representations of the Galilean group, and use this information to build models of interacting systems of particles and fields.

The methods of finding unitary representations for the Galilean group in non-relativistic mechanics is similar to the same problem for the Poincaré group in relativistic mechanics. The results for the Poincaré group are, perhaps, better known to physicists and well described in Weinberg[**?**, Chapter 2], for example. It turns out, however, that the group structure of the Galilean group is not not as simple as that of the Poincaré group. The landmark paper by Bargmann[2] on unitary projective representations of continuous groups contains theorems and results which we use here. Ray representations of the Galilean group are also discusses by Hamermesh[**?**][p. 484]. We also use results from several papers by Levy-Leblond[3, 4, 5, 6] on the Galilei group. In the next section, we show that Galilean transformation form a group.

9.1 Galilean transformations

A Galilean transformation includes time and space translation, space rotations, and velocity boosts of the coordinate system. An "event" in a coordinate frame Σ is given by the coordinates $(\mathbf{x}, t)$. The same event is described by the coordinates $(\mathbf{x}', t')$ in another frame Σ', which is rotated an amount R, displaced a distance $\mathbf{a}$, moving at a velocity $\mathbf{v}$, and using a clock running at a time $t' = t + \tau$, with respect to frame Σ, as shown in Fig. 9.1. The relation between the events in Σ and Σ' is given by the proper Galilean transformation:

$$\mathbf{x}' = R(\mathbf{x}) + \mathbf{v}t + \mathbf{a}\,, \qquad t' = t + \tau\,, \tag{9.3}$$

with R a proper real three-dimensional orthogonal matrix such that $\det R = +1$. We regard the transformation (9.3) as a relationship between an event *as viewed from two different coordinate frames*. The basic premise of non-relativistic quantum mechanics of point particles is that it is impossible to distinguish between these two coordinate systems and so this space-time symmetry must be a property of the vector space which describes the physical system. We discuss *improper* transformations in Section 9.7.

9.1.1 The Galilean group

We need to show that elements of a Galilean transformation form a group. We write the transformation as: $\Sigma' = G(\Sigma)$, where Σ refers to the coordinate system and $G = (R, \mathbf{v}, \mathbf{a}, \tau)$ to the elements describing the transformation. A group of elements is defined by the following four requirements:

Definition 17 (group)**.** A group $\mathcal{G}$ is a set of objects, the elements of the group, which we call G, and a multiplication, or combination, rule for combining any two of them to form a product, subject to the following four conditions:

1. The product $G_1 G_2$ of any two group elements must be another group element G_3.
2. Group multiplication is associative: $(G_1 G_2) G_3 = G_1 (G_2 G_3)$.
3. There is a unique group element I, called the identity, such that $I\,G = G$ for all G in the group.
4. For any G there is an inverse, written G^{-1} such that $G\,G^{-1} = G^{-1}\,G = I$.

We first show that one Galilean transformation followed by a second Galilean transformation is also a Galilean transformation. This statement is contained in the following theorem:

Theorem 17 (Composition rule)**.** *The multiplication law for the Galilean group is*

$$\begin{aligned} G'' = G'G &= (R', \mathbf{v}', \mathbf{a}', \tau')\,(R, \mathbf{v}, \mathbf{a}, \tau)\,, \\ &= (R'R, \mathbf{v}' + R'\mathbf{v}, \mathbf{a}' + R'\mathbf{a} + \mathbf{v}'\tau, \tau' + \tau)\,. \end{aligned} \tag{9.4}$$

Proof. We find:

$$\begin{aligned} \mathbf{x}' &= R\mathbf{x} + \mathbf{v}t + \mathbf{a}\,, \qquad t' = t + \tau\,, \\ \mathbf{x}'' &= R'\mathbf{x}' + \mathbf{v}'t' + \mathbf{a}' = R'R\,\mathbf{x} + (R'\mathbf{v} + \mathbf{v}')t + R'\mathbf{a} + \mathbf{v}'\tau + \mathbf{a}' \\ &\equiv R''\mathbf{x} + \mathbf{v}''t + \mathbf{a}'' \\ t'' &= t' + \tau' = t + \tau + \tau' \equiv t + \tau'' \end{aligned}$$

where

$$\begin{aligned} R'' &= R'R\,, & \mathbf{v}'' &= R'\mathbf{v} + \mathbf{v}' \\ \mathbf{a}'' &= R'\mathbf{a} + \mathbf{v}'\tau + \mathbf{a}' & \tau'' &= \tau' + \tau\,. \end{aligned}$$

That is, R'' is also an orthogonal matrix with unit determinant, and $\mathbf{v}''$ and $\mathbf{a}''$ are vectors. □

Thus the Galilean group $\mathcal{G}$ is the set of all elements $G = (R, \mathbf{v}, \mathbf{a}, \tau)$, consisting of ten real parameters, three for the rotation matrix R, three each for boosts $\mathbf{v}$ and for space translations $\mathbf{a}$, and one for time translations τ.

Definition 18. The identity element is $1 = (1, 0, 0, 0)$, and the inverse element of G is:

$$G^{-1} = (R^{-1}, -R^{-1}\mathbf{v}, -R^{-1}(\mathbf{a} - \mathbf{v}\tau), -\tau)\,, \tag{9.5}$$

as can be easily checked.

Thus the elements of Galilean transformations form a group.

Example 26 (Matrix representation)**.** It is easy to show that the following 5×5 matrix representation of the Galilean group elements:

$$\underline{G} = \begin{pmatrix} R & \mathbf{v} & \mathbf{a} \\ 0 & 1 & \tau \\ 0 & 0 & 1 \end{pmatrix}\,, \tag{9.6}$$

forms a group, where group multiplication is defined to be *matrix* multiplication: $\underline{G}'' = \underline{G}'\underline{G}$. Here R is understood to be a 3×3 matrix and $\mathbf{v}$ and $\mathbf{a}$ are 3×1 column vectors.

Remark 14. An infinitesimal Galilean transformation *of the coordinate system* is given in vector notation by:

$$\begin{aligned} \Delta\mathbf{x} &= \Delta\theta\,\mathbf{x}\times\hat{\mathbf{n}} + \Delta\mathbf{v}\,t + \Delta\mathbf{a}\,, \\ \Delta t &= \Delta\tau\,. \end{aligned} \tag{9.7}$$

The elements of the transformation are given by $1+\Delta G$, where $\Delta G = (\,\Delta\theta, \Delta\mathbf{v}, \Delta\mathbf{a}, \Delta\tau\,)$.

Example 27. We can find differential representations of the generators of the transformation in classical physics. We start by considering complex functions $\psi(\mathbf{x},t)$ which transform "like scalars" under Galilean transformations, that is:

$$\psi'(\mathbf{x}',t') = \psi(\mathbf{x},t)\,. \tag{9.8}$$

For infinitesimal transformations, this reads:

$$\psi'(\mathbf{x}',t') = \psi(\mathbf{x}'-\Delta\mathbf{x}, t'-\delta t) = \psi(\mathbf{x}',t') - \Delta\mathbf{x}\cdot\boldsymbol{\nabla}'\,\psi(\mathbf{x}',t') - \Delta t\,\partial_{t'}\,\psi(\mathbf{x}',t') + \cdots\,, \tag{9.9}$$

and, to first order, the *change in functional form* of $\psi(\mathbf{x},t)$ is given by:

$$\Delta\psi(\mathbf{x},t) = -\big\{\,\Delta\mathbf{x}\cdot\boldsymbol{\nabla} + \Delta t\,\partial_t\,\big\}\,\psi(\mathbf{x},t)\,, \tag{9.10}$$

Here we have put $\mathbf{x}'\to\mathbf{x}$ and $t'\to t$. Substituting (9.7) into the above gives:

$$\Delta\psi(\mathbf{x},t) = -\big\{\,-\Delta\theta\,\hat{\mathbf{n}}\cdot\mathbf{x}\times\boldsymbol{\nabla} + t\,\Delta\mathbf{v}\cdot\boldsymbol{\nabla} + \Delta\mathbf{a}\cdot\boldsymbol{\nabla} + \Delta\tau\,\partial_t\,\big\}\,\psi(\mathbf{x},t)\,. \tag{9.11}$$

We define the ten differential **generator** operators $(\mathbf{J},\mathbf{K},\mathbf{P},H)$ of Galilean transformations by

$$\Delta\psi(\mathbf{x},t) = \frac{i}{\hbar}\,\big\{\,\Delta\theta\,\hat{\mathbf{n}}\cdot\mathbf{J} + \Delta\mathbf{v}\cdot\mathbf{K} - \Delta\mathbf{a}\cdot\mathbf{P} + \Delta\tau\,H\big\}\,\psi(\mathbf{x},t)\,, \tag{9.12}$$

Here we have introduced a constant $\hbar$ so as to make the units of $\mathbf{J}$, $\mathbf{K}$, $\mathbf{P}$, and H to be the classical units of angular momentum, impulse, linear momentum, and energy, respectively.[1] Comparing (9.11) to (9.12), we find classical differential representations of the generators:

$$\mathbf{J} = \frac{\hbar}{i}\,\mathbf{x}\times\boldsymbol{\nabla}\,,\qquad \mathbf{K} = -\frac{\hbar t}{i}\,\boldsymbol{\nabla}\,,\qquad \mathbf{P} = \frac{\hbar}{i}\,\boldsymbol{\nabla}\,,\qquad H = i\hbar\,\frac{\partial}{\partial t}\,. \tag{9.13}$$

When acting on complex functions $\psi(\mathbf{x},t)$, these ten generators produce the corresponding changes in the functional form of the functions.

Example 28. Using the differential representation (9.13), it is easy to show that the generators obey the algebra:

$$\begin{aligned} &[J_i,J_j] = i\hbar\,\epsilon_{ijk}J_k\,, && [K_i,K_j] = 0\,, && [J_i,H] = 0\,, \\ &[J_i,K_j] = i\hbar\,\epsilon_{ijk}K_k\,, && [P_i,P_j] = 0\,, && [P_i,H] = 0\,, \\ &[J_i,P_j] = i\hbar\,\epsilon_{ijk}P_k\,, && [K_i,P_j] = 0\,, && [K_i,H] = i\hbar\;P_i\,. \end{aligned} \tag{9.14}$$

9.1.2 Group structure

If the generators of a group all commute, then the group is called **Abelian**. An invariant Abelian subgroup consists of a subset of generators that commute with each other and whose commutators with any other member of the group also belong to the subgroup. For the Galilean group, the largest Abelian subgroup is the six-parameter group $\mathcal{U} = [\mathbf{L},\mathbf{P}]$ generating boosts and translations. The largest abelian subgroup of the **factor group**, $\mathcal{G}/\mathcal{U}$, is the group $\mathcal{D} = [H]$, generating time translations. This leaves the **semi-simple** group

[1] The *size* of $\hbar$ is fixed by the physics.

$\mathcal{R} = [\mathbf{J}]$, generating rotations. A semi-simple group is one which transform among themselves and cannot be reduced further by removal of an Abelian subgroup. So the Galilean group can be written as the semidirect product of a six parameter abelian group $\mathcal{U}$ with the semidirect product of a one parameter abelian group $\mathcal{D}$ by a three parameter simple group $\mathcal{R}$,

$$\mathcal{G} = (\mathcal{R} \times \mathcal{D}) \times \mathcal{U} \,. \tag{9.15}$$

In contrast, the Poincaré group is the simidirect product of a simple group $\mathcal{L}$ generating Lorentz transformations by an abelian group $\mathcal{C}$ generating space and time translations,

$$\mathcal{P} = \mathcal{L} \times \mathcal{C} \,. \tag{9.16}$$

9.2 Galilean transformations in quantum mechanics

Now let $|\,\psi(\Sigma)\,\rangle$ be a vector in $\mathcal{V}$ which refers to a specific coordinate system Σ and let $|\,\psi(\Sigma')\,\rangle$ be a vector which refers to the coordinate system $\Sigma' = G\Sigma$. Then we know by Wigner's theorem that:

$$|\,\psi(\Sigma')\,\rangle = U(G)\,|\,\psi(\Sigma)\,\rangle \,, \tag{9.17}$$

where $U(G)$ is unitary.[2] In non-relativistic quantum mechanics, we want to find unitary transformations $U(G)$ for the Galilean group. We do this by applying the classical group multiplication properties to unitary transformations. That is, if (9.17) represents a transformation from Σ to Σ' by G, and a similar relation holds for a transformation from Σ' to Σ'' by G', then the combined transformation is given by:

$$|\,\psi(\Sigma'')\,\rangle = U(G')\,|\,\psi(\Sigma')\,\rangle = U(G')\,U(G)\,|\,\psi(\Sigma)\,\rangle \,. \tag{9.18}$$

However the direct transformation from Σ to Σ'' is given classically by $G'' = G'G$, and quantum mechanically by:

$$|\,\psi(\Sigma'')\,\rangle' = U(G'')\,|\,\psi(\Sigma)\,\rangle = U(G'G)\,|\,\psi(\Sigma)\,\rangle \,. \tag{9.19}$$

Now $|\,\psi(\Sigma'')\,\rangle$ and $|\,\psi(\Sigma'')\,\rangle'$ *must belong to the same ray,* and therefore can only differ by a phase. Thus we can deduce that:

$$U(G')\,U(G) = e^{i\phi(G',G)/\hbar}\,U(G'G) \,, \tag{9.20}$$

where $\phi(G', G)$ is real and depends only on the group elements G and G'. Unitary representations of operators which obey Eq. (9.20) with non-zero phases are called *projective* representations. If the phase $\phi(G', G) = 0$, they are called *faithful* representations. The Galilean group generally is projective, not faithful.[3] The group composition rule, Eq. (9.20), will be used to find the unitary transformation $U(G)$.

Now we can take the unit element to be: $U(1) = 1$. So using the group composition rule (9.20), unitarity requires that:

$$U^\dagger(G)U(G) = U^{-1}(G)U(G) = U(G^{-1})U(G) = e^{i\phi(G^{-1},G)/\hbar}\,U(1,0) = 1 \,. \tag{9.21}$$

so that $\phi(G^{-1}, G) = 0$. We will use this unitarity requirement in section 9.2.1 below.

Infinitesimal transformations are generated from the unity element by the set $\Delta G = (\Delta\omega, \Delta\mathbf{v}, \Delta\mathbf{a}, \Delta\tau)$, where $\Delta\omega_{ij} = \epsilon_{ijk}n_k\Delta\theta = -\Delta\omega_{ji}$ is an antisymmetric matrix. We write the unitary transformation for this infinitesimal transformation as:

$$\begin{aligned} U(1 + \Delta G) &= 1 + \frac{i}{\hbar}\left\{\Delta\omega_{ij}\,J_{ij}/2 + \Delta v_i\,K_i - \Delta a_i\,P_i + \Delta\tau\,H\right\} + \cdots \\ &= 1 + \frac{i}{\hbar}\left\{\Delta\theta\,\hat{\mathbf{n}}\cdot\mathbf{J} + \Delta\mathbf{v}\cdot\mathbf{K} - \Delta\mathbf{a}\cdot\mathbf{P} + \Delta\tau\,H\right\} + \cdots \,, \end{aligned} \tag{9.22}$$

[2] We will consider anti-unitary symmetry transformations later.

[3] In contrast, the Poincaré group *is* faithful.

where J_i, K_i, P_i and H are operators on $\mathcal{V}$ which generater rotations, boosts, and space and time translations, respectively. Here $\Delta\omega_{ij} = \epsilon_{ijk}\, n_k\, \Delta\theta$ is an antisymmetric matrix representing an infinitesimal rotation about an axis defined by the unit vector n_k by an angle $\Delta\theta$. In a similar way, we write the antisymmetric matrix of operators J_{ij} as $J_{ij} = \epsilon_{ijk}J_k$, where J_k is a set of three operators.

Remark 15. Again, we have introduced a constant $\hbar$ so that the units of the operators $\mathbf{J}$, $\mathbf{K}$, $\mathbf{P}$, and H are given by units of angular momentum, impulse, linear momentum, and energy, respectively. The value of $\hbar$ must be fixed by experiment.[4]

Remark 16. The sign of the operators P_i and H, relative to J_k in (9.22) is arbitrary — the one we have chosen is conventional.

In the next section, we find the phase factor $\phi(G';G)$ in Eq. (9.20) for unitary representations of the Galilean group.

9.2.1 Phase factors for the Galilean group.

The phases $\phi(G', G)$ must obey basic properties required by the transformation rules. Since $U^{-1}(G)U(G) = U(G^{-1})U(G) = 1$, we find from the unitarity requirement (9.21),

$$\phi(G^{-1}, G) = 0\,. \tag{9.23}$$

Also, the associative law for group transformations,

$$U(G'')\,(U(G')U(G)) = (U(G'')U(G'))\,U(G)\,,$$

requires that

$$\phi(G'', G'G) + \phi(G', G) = \phi(G'', G') + \phi(G''G', G)\,. \tag{9.24}$$

From (9.23) and (9.24), we easily obtain $\phi(1,1) = \phi(1,G) = \phi(G,1) = 0$. Eqs. (9.23) and (9.24) are the *defining equations* for the phase factor $\phi(G', G)$, and will be used in Bargmann's theorem (18) to find the phase factor below.

Note that (9.23) and (9.24) can be satisfied by any $\phi(G', G)$ of the form

$$\phi(G', G) = \chi(G'G) - \chi(G') - \chi(G)\,. \tag{9.25}$$

Then the phase can be eliminated by a trivial change of phase of the unitary transformation, $\bar{U}(G) = e^{i\chi(G)}U(G)$. Thus two phases $\phi(G', G)$ and $\phi'(G', G)$ which differ from each other by functions of the form (9.25) are equivalent. For Galilean transformations, unlike the case for the Poincaré group, the phase $\phi(G', G)$ *cannot* be eliminated by a simple redefinition of the unitary operators. This phase makes the stude of unitary representations of the Galilean group much harder than the Poincaré group in relativistic quantum mechanics.

It turns out that the phase factors for the Galilean group are not easy to find. The result is stated in a theorem due to Bargmann[2]:

Theorem 18 (Bargmann)**.** *The phase factor for the Galilean group is given by:*

$$\phi(G', G) = \frac{M}{2}\{\,\mathbf{v}' \cdot R'(\mathbf{a}) - \mathbf{v}' \cdot R'(\mathbf{v})\,\tau - \mathbf{a}' \cdot R'(\mathbf{v})\,\}\,, \tag{9.26}$$

with M any real number.

[4]Plank introduced $\hbar$ in order to make the classical partition function dimensionless. The value of $\hbar$ was fixed by the experimental black-body radiation law.

Proof. A proper Galilean transformation is given by Eq. (9.3). The group multiplication rules are given in Eq. (9.4):

$$\begin{aligned} R'' &= R'R\,, \\ \mathbf{v}'' &= \mathbf{v}' + R'(\mathbf{v})\,, \\ \mathbf{a}'' &= \mathbf{a}' + \mathbf{v}'\tau + R'(\mathbf{a})\,, \\ \tau'' &= \tau' + \tau\,. \end{aligned} \tag{9.27}$$

We first note that $\mathbf{v}$ and $\mathbf{a}$ *transform linearly*. Therefore, it is useful to introduce a six-component column matrix ξ and a 6×6 matrix $\Theta(\tau)$, which we write as:

$$\xi = \begin{pmatrix} \mathbf{v} \\ \mathbf{a} \end{pmatrix}\,, \qquad \Theta(\tau) = \begin{pmatrix} 1 & 0 \\ \tau & 1 \end{pmatrix}\,, \tag{9.28}$$

so that we can write the group multiplication rules for these parameters as:

$$\xi'' = \Theta(\tau)\,\xi' + R'\,\xi\,, \tag{9.29}$$

which is linear in the ξ variables. We label the rest of the parameters by $g = (R,\tau)$, which obey the group multiplication rules:

$$R'' = R'R\,, \qquad \tau'' = \tau' + \tau\,. \tag{9.30}$$

We note here that the unit element of g is $g = (1,0)$. We also note that the matrices $\Theta(\tau)$ are a faithful representation of the subgroup of τ transformations. That is, we find:

$$\Theta(\tau'') = \Theta(\tau')\,\Theta(\tau)\,. \tag{9.31}$$

We seek now the form of $\phi(G',G)$ by solving the defining equation (9.24):

$$\phi(G'',G'G) + \phi(G',G) = \phi(G'',G') + \phi(G''G',G)\,. \tag{9.32}$$

The only way this can be satisfied is if $\phi(G',G)$ is bilinear in ξ, because the transformation of these variables is linear. Thus we make the Ansatz:

$$\phi(G',G) = \xi'^{\,T}\,\Phi(g',g)\,\xi\,, \tag{9.33}$$

where $\Phi(g',g)$ is a 6×6 matrix, but depends only on the elements g and g'. We now work out all four terms in Eq. (9.32). We find:

$$\begin{aligned} \phi(G'',G'G) &= \xi''^{\,T}\,\Phi(g'',g'g)\,\big[\,\Theta(\tau)\,\xi' + R'\,\xi\,\big] \\ &= \xi''^{\,T}\,\Phi(g'',g'g)\,\Theta(\tau)\,\xi' + \xi''^{\,T}\,\Phi(g'',gg)\,R'\,\xi\,, \\ \phi(G',G) &= \xi'^{\,T}\,\Phi(g',g)\,\xi\,, \\ \phi(G'',G') &= \xi''^{\,T}\,\Phi(g'',g')\,\xi'\,, \\ \phi(G''G',G) &= \big[\,\xi'^{T}R''^{T} + \xi''^{T}\Theta^{T}(\tau')\,\big]\,\Phi(g''g',g)\,\xi \\ &= \xi'^{T}R''^{T}\,\Phi(g''g',g)\,\xi + \xi''^{T}\Theta^{T}(\tau')\,\Phi(g''g',g)\,\xi\,. \end{aligned} \tag{9.34}$$

Substituting these results into (9.32), and equating coefficients for the three bilinear forms, we find for the three pairs: $(\xi';\xi)$, $(\xi'';\xi')$, and $(\xi'';\xi)$:

$$\Phi(g',g) = R''^{T}\,\Phi(g''g',g)\,, \tag{9.35}$$

$$\Phi(g'',g'g)\,\Theta(\tau) = \Phi(g'',g') \tag{9.36}$$

$$\Phi(g'',g'g)\,R' = \Theta^{T}(\tau')\,\Phi(g''g',g)\,. \tag{9.37}$$

These relations provide functional equations for the matrix elements. We start by using the orthogonality of R and writing (9.35) in the form:

$$\Phi(g''g', g) = R''\,\Phi(g', g) \tag{9.38}$$

Since g' is arbitrary, we can set it equal the unit element: $g' = (1, 0)$. Then $g''g' = g''$, and we find:

$$\Phi(g'', g) = R''\,\Phi(1, g)\,. \tag{9.39}$$

When this result is substituted into (9.36) and (9.37), we find:

$$R''\,\Phi(1, g'g)\,\Theta(\tau) = R''\,\Phi(1, g') \tag{9.40}$$

$$R''\,\Phi(1, g'g)\,R' = \Theta^T(\tau')\,R''R'\,\Phi(1, g)\,. \tag{9.41}$$

and from (9.40), we find:

$$\Phi(1, g'g)\,\Theta(\tau) = \Phi(1, g')\,. \tag{9.42}$$

Here g' is arbitrary, so that we can it to the unit element: $g' = 1$, and find:

$$\Phi(1, g)\,\Theta(\tau) = \Phi(1, 1)\,. \tag{9.43}$$

Now in (9.41), R'' and R' act only on vectors and commute with the matrices Θ and Φ, so we can write this as:

$$\Phi(1, g'g) = \Theta^T(\tau')\,\Phi(1, g)\,. \tag{9.44}$$

Again in (9.44), we can set $g = 1$, from which we find:

$$\Phi(1, g') = \Theta^T(\tau')\,\Phi(1, 1)\,. \tag{9.45}$$

So combining (9.43) and (9.45), we find that $\Phi(1, 1)$ must satisfy the equation:

$$\Phi(1, 1) = \Theta^T(\tau)\,\Phi(1, 1)\,\Theta(\tau)\,, \tag{9.46}$$

for all values of τ. Which means that $\Phi(1, 1)$ must be a constant 6×6 matrix, independent of τ. In order to solve (9.46), we write out $\Phi(1, 1)$ in component form:

$$\Phi(1, 1) = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}\,, \tag{9.47}$$

so that (9.46) requires:

$$\Phi_{11} = \Phi_{11} + \tau\,(\Phi_{12} + \Phi_{21}) + \tau^2\,\Phi_{22}\,, \tag{9.48}$$

$$\Phi_{12} = \Phi_{12} + \tau\,\Phi_{22}\,, \tag{9.49}$$

$$\Phi_{21} = \Phi_{21} + \tau\,\Phi_{22}\,, \tag{9.50}$$

$$\Phi_{22} = \Phi_{22}\,, \tag{9.51}$$

which must hold for all values of τ. This is possible only if $\Phi_{22} = 0$, and that $\Phi_{21} = -\Phi_{12}$. Φ_{11} is then arbitrary. So let us put $\Phi_{12} = M/2$ and $\Phi_{11} = M'/2$. So the general solution for the phase matrix contains two constants. We write the result as:

$$\Phi(1, 1) = \frac{M}{2}\,Z + \frac{M'}{2}Z'\,, \qquad \text{where} \qquad Z = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\,, \qquad Z' = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\,, \tag{9.52}$$

From Eqs. (9.33), (9.39), and (9.45), we find:

$$\phi(G', G) = \xi'^T\,\Phi(g', g)\,\xi\,, \qquad \Phi(g', g) = \Theta^T(\tau)\,\Phi(1, 1)\,R'\,. \tag{9.53}$$

Recall that R' commutes with $\Theta(\tau)$ and $\Phi(1,1)$. It turns out that the term involving $M'Z'$ is a trivial phase. For this term, we find:

$$\begin{aligned}\phi_{Z'}(G',G) &= \frac{M'}{2}\xi'^{\,T}\,\Theta^T(\tau)\,Z'\,R'(\xi)\\ &= \frac{M'}{2}\,\mathbf{v}'\cdot R'(\mathbf{v}) = \frac{M'}{4}\left\{\,v''^{\,2}-v^2-v'^{\,2}\,\right\},\end{aligned} \tag{9.54}$$

So (9.54) is a trivial phase and can be absorbed into the definition of $U(g)$. So then from Eq. (9.52), the phase is given by:

$$\begin{aligned}\phi(G',G) &= +\frac{M}{2}\,\xi'^{\,T}\,\Theta^T(\tau)\,Z\,R'(\xi) = -\frac{M}{2}\,[\,R'(\xi)\,]^T\,Z\,\Theta(\tau)\,\xi'\,.\\ &= \frac{M}{2}\{\,\mathbf{v}'\cdot R'(\mathbf{a})-\mathbf{v}'\cdot R'(\mathbf{v})\,\tau-\mathbf{a}'\cdot R'(\mathbf{v})\,\}\,,\end{aligned} \tag{9.55}$$

which is what we quoted in the theorem. In the first line, we have used the fact that Z is antisymmetric: $Z^T=-Z$. This phase is non-trival! For example, we might try to do the same tricks we used for the trival phase in Eq. (9.54), and write:

$$\begin{aligned}\xi''^{\,T}\,Z\,\xi'' &= \left\{\,[\,R'(\xi)\,]^T+\xi'^{\,T}\Theta^T(\tau)\,\right\}Z\left\{\,\Theta(\tau)\,\xi'+R'(\xi)\,\right\}\\ &= \xi'^{\,T}\,Z\,\xi'+\xi^T\,Z\,\xi+\xi'^{\,T}\,\Theta^T(\tau)\,Z\,R'(\xi)+[\,R'(\xi)\,]^T\,Z\,\Theta(\tau)\,\xi'\,.\end{aligned} \tag{9.56}$$

But the last two terms *cancel* rather than add because of the antisymmetry of Z. So we cannot turn (9.55) into a trival phase the way we did for (9.54). This completes the proof. □

Remark 17. Bargmann gave this phase in his classic paper on continuous groups[2], and indicated how he found it in a footnote to that paper. Notice that M appears here as an undetermined multiplicative parameter. Since we have introduced a constant $\hbar$ with the dimensions of action in the definition of the phase, *M has units of mass.*

We can write the phase as:

$$\phi(G',G) = \tfrac{1}{2}M\,R'_{ij}\{v'_i a_j - a'_i v_j - v'_i v_j \tau] \tag{9.57}$$

Notice that $\phi(G^{-1},G)=0$.

The phase for infinitesimal transformations are given by:

$$\begin{aligned}\phi(G,1+\Delta G) &= \tfrac{1}{2}M\,R_{ij}[v_i\Delta a_j - a_i\Delta v_j]+\cdots\,,\\ \phi(1+\Delta G,G) &= \tfrac{1}{2}M\,[\Delta v_i(a_i-v_i\tau)-\Delta a_i v_i]+\cdots\,,\end{aligned} \tag{9.58}$$

Next, we find the transformation properties of the generators.

9.2.2 Unitary transformations of the generators

In this section, we find the unitary transformation $U(G)$ for the generators of the Galilean group. We start by finding the transformation rules for all the generators. This is stated in the following theorem:

Theorem 19. *The generators transform according to the rules:*

$$U^\dagger(G)\,\mathbf{J}\,U(G) = R\{\mathbf{J}+\mathbf{K}\times\bar{\mathbf{v}}+\bar{\mathbf{a}}\times(\mathbf{P}+M\,\bar{\mathbf{v}})\}\,, \tag{9.59}$$

$$U^\dagger(G)\,\mathbf{K}\,U(G) = R\{\mathbf{K}-(\mathbf{P}+M\bar{\mathbf{v}})\,\tau+M\bar{\mathbf{a}}\}\,, \tag{9.60}$$

$$U^\dagger(G)\,\mathbf{P}\,U(G) = R\{\mathbf{P}+M\bar{\mathbf{v}}\}\,, \tag{9.61}$$

$$U^\dagger(G)\,H\,U(G) = H+\bar{\mathbf{v}}\cdot\mathbf{P}+\tfrac{1}{2}Mv^2\,. \tag{9.62}$$

where $\bar{\mathbf{v}}=R^{-1}(\mathbf{v})$ *and* $\bar{\mathbf{a}}=R^{-1}(\mathbf{a})$.

Proof. We start by considering the transformations:

$$U^\dagger(G)\,U(1+\Delta G)\,U(G)\,, \tag{9.63}$$

where G and $1+\Delta G$ are two *different* transformations. On one hand, using the definition (9.22) for infinitesimal transformations in terms of the generators, (9.63) is given by:

$$\begin{aligned} 1+\frac{i}{2\hbar}\,\Delta\omega_{ij}\,U^\dagger(G)\,J_{ij}\,U(G)+\frac{i}{\hbar}\,\Delta v_i\,U^\dagger(G)\,K_i\,U(G) \\ -\frac{i}{\hbar}\,\Delta a_i\,U^\dagger(G)\,P_i\,U(G)+\frac{i}{\hbar}\,\Delta\tau\,U^\dagger(G)\,H\,U(G)+\cdots \end{aligned} \tag{9.64}$$

On the other, using the composition rule (9.20), Eq. (9.63) can be written as:

$$\begin{aligned} e^{i[\phi(G^{-1},(1+\Delta G)G)+\phi((1+\Delta G),G)]/\hbar}\,U(G^{-1}(1+\Delta G)G) \\ =e^{i\chi(G,\Delta G)/\hbar}\,U(1+\Delta G')\,. \end{aligned} \tag{9.65}$$

where $\Delta G'=G^{-1}\,\Delta G\,G$. Working out this transformation, we find the result:

$$\begin{aligned} \Delta\omega'_{ij}&=R_{ki}R_{lj}\,\Delta\omega_{kl}\,,\\ \Delta v'_i&=R_{ji}\,(\Delta\omega_{jk}\,v_k+\Delta v_j)\,,\\ \Delta a'_i&=R_{ji}\,(\Delta\omega_{jk}\,a_k+\Delta v_j\tau+\Delta a_j-v_j\Delta\tau)\\ \Delta\tau'&=\Delta\tau\,, \end{aligned}$$

and the phase $\chi(G,\Delta G)$ is defined by:

$$\chi(G,\Delta G)=\phi(G^{-1},(1+\Delta G)G)+\phi(1+\Delta G,G)\,. \tag{9.66}$$

We can simplify the calculation of the phase using an identity derived from (9.24):

$$\begin{aligned} \phi(G,G^{-1}(1+\Delta G)G)+\phi(G^{-1},(1+\Delta G)G)\\ =\phi(G,G^{-1})+\phi(GG^{-1},(1+\Delta G)G)=\phi(1,(1+\Delta G)G)=0\,, \end{aligned}$$

and therefore, since $G^{-1}(1+\Delta G)G=1+\Delta G'$, we have:

$$\phi(G^{-1},(1+\Delta G)G)=-\phi(G,1+\Delta G')\,.$$

So the phase $\chi(G,\Delta G)$ is given by:

$$\chi(G,\Delta G)=\phi(1+\Delta G,G)-\phi(G,1+\Delta G')\,. \tag{9.67}$$

Now using (9.58), we find to first order:

$$\begin{aligned} \phi(1+\Delta G,G)&=\tfrac{1}{2}M\,[\Delta v_i(a_i-v_i\tau)-\Delta a_i v_i]+\cdots\,,\\ \phi(G,1+\Delta G')&=\tfrac{1}{2}M\,R_{ij}[v_i\Delta a'_j-a_i\Delta v'_j]+\cdots\,,\\ &=\tfrac{1}{2}M\,\{v_i(\Delta\omega_{ij}a_j+\Delta v_i\tau+\Delta a_i-v^2\Delta\tau)-a_i(\Delta v_i+\Delta\omega_{ij}v_j)\}\\ &\quad+\cdots\,, \end{aligned}$$

from which we find,

$$\begin{aligned} \chi(G,\Delta G)=\tfrac{1}{2}\Delta\omega_{ij}\,M(a_iv_j-a_jv_i)+\Delta v_i\,M(a_i-v_i\tau)-\Delta a_i\,Mv_i\\ +\Delta\tau\,\tfrac{1}{2}Mv^2+\cdots\,. \end{aligned} \tag{9.68}$$

For the unitary operator $U(1+\Delta G')$, we find:

$$\begin{aligned} U(1+\Delta G') &= 1 + \frac{i}{2\hbar}\,\Delta\omega'_{ij}\,J_{ij} + \frac{i}{\hbar}\,\Delta v'_i\,K_i - \frac{i}{\hbar}\,\Delta a'_i\,P_i + \frac{i}{\hbar}\,\Delta\tau'\,H + \cdots\,, \\ &= 1 + \frac{i}{2\hbar}\,\Delta\omega_{ij}\,[R_{ik}R_{jl}J_{kl} + 2R_{il}(v_jK_l - a_jP_l)] \\ &\quad + \frac{i}{\hbar}\,\Delta v_i\,R_{ij}(K_j - \tau P_j) - \frac{i}{\hbar}\,\Delta a_i\,R_{ij}P_j \\ &\quad + \frac{i}{\hbar}\,\Delta\tau\,(H + R_{ij}v_iP_j) + \cdots\,, \end{aligned} \tag{9.69}$$

Combining relations (9.68) and (9.69), we find, to first order, the expansion:

$$\begin{aligned} &e^{i\chi(G,\Delta G)/\hbar}\,U(1+\Delta G') \\ &\quad = 1 + \frac{i}{2\hbar}\,\Delta\omega_{ij}\,[R_{ik}R_{jl}J_{kl} + 2R_{il}(v_jK_l - a_jP_l) + M(a_iv_j - a_jv_i)] \\ &\qquad + \frac{i}{\hbar}\,\Delta v_i\,[R_{ij}(K_j - \tau P_j) + M(a_i - v_i\tau)] \\ &\qquad - \frac{i}{\hbar}\,\Delta a_i\,[R_{ij}P_j + Mv_i] \\ &\qquad + \frac{i}{\hbar}\,\Delta\tau\,[H + R_{ij}v_iP_j + \tfrac{1}{2}Mv^2] + \cdots\,, \end{aligned} \tag{9.70}$$

Comparing coefficients of $\Delta\omega_{ij}$, Δv_i, Δa_i, and $\Delta\tau$ in (9.64) and (9.70), we get:

$$\begin{aligned} U^\dagger(G)\,J_{ij}\,U(G) &= R_{ik}R_{jl}J_{kl} + 2R_{il}(v_jK_l - a_jP_l) + M(a_iv_j - a_jv_i) \\ &= R_{ik}R_{jl}J_{kl} + (K'_iv_j - K'_jv_i) - (P'_ia_j - P'_ja_i) + M(a_iv_j - a_jv_i) \\ U^\dagger(G)\,K_i\,U(G) &= R_{ij}(K_j - \tau P_j) + M(a_i - v_i\tau) \\ U^\dagger(G)\,P_i\,U(G) &= R_{ij}P_j + Mv_i \\ U^\dagger(G)\,H\,U(G) &= H + v_iP'_i + \tfrac{1}{2}Mv^2 \end{aligned}$$

where, $K'_i = R_{ij}K_j$ and $P'_i = R_{ij}P_j$. In the second line, we have used the antisymmetry of J_{ij}. These equations simplify if we rewrite them in terms of the components of the angular momentum vector J_k rather than the antisymmetric tensor J_{ij}. We have the definitions:

$$\begin{aligned} J_{ij} &= \epsilon_{ijk}J_k\,, \\ K'_iv_j - K'_jv_i &= \epsilon_{ijk}[\mathbf{K}'\times\mathbf{v}]_k\,, \\ P'_ia_j - P'_ja_i &= \epsilon_{ijk}[\mathbf{P}'\times\mathbf{a}]_k\,, \\ v_ia_j - v_ja_i &= \epsilon_{ijk}[\mathbf{v}\times\mathbf{a}]_k\,. \end{aligned}$$

The identity,

$$R_{ik}R_{jl}\,\epsilon_{klm} = \det[\,R\,]\,\epsilon_{ijn}R_{nm}\,, \tag{9.71}$$

is obtained from the definition of the determinant of R and the orthogonality relations for R. For proper transformations, which is what we consider here, $\det[\,R\,] = 1$. So the above equations become, in vector notation,

$$\begin{aligned} U^\dagger(G)\,\mathbf{J}\,U(G) &= R\{\mathbf{J} + \mathbf{K}\times\bar{\mathbf{v}} + \bar{\mathbf{a}}\times(\mathbf{P} + M\,\bar{\mathbf{v}})\}\,, \\ U^\dagger(G)\,\mathbf{K}\,U(G) &= R\{\mathbf{K} - (\mathbf{P} + M\bar{\mathbf{v}})\,\tau + M\bar{\mathbf{a}}\}\,, \\ U^\dagger(G)\,\mathbf{P}\,U(G) &= R\{\mathbf{P} + M\bar{\mathbf{v}}\}\,, \\ U^\dagger(G)\,H\,U(G) &= H + \bar{\mathbf{v}}\cdot\mathbf{P} + \tfrac{1}{2}Mv^2\,. \end{aligned}$$

where $\bar{\mathbf{v}} = R^{-1}(\mathbf{v})$ and $\bar{\mathbf{a}} = R^{-1}(\mathbf{a})$. This completes the proof of the theorem, as stated. □

Exercise 8. Using the indentity (9.71) with $\det[\,R\,] = +1$, show that $R(\mathbf{A}\times\mathbf{B}) = R(\mathbf{A})\times R(\mathbf{B})$.

We next turn to a discussion of the commutation relations for the generators.

9.2.3 Commutation relations of the generators

In this section, we prove a theorem which gives the commutation relations for the generators of the Galilean group. The set of commutation relations for the group can be thought of as rules for "multiplying" any two operators, and are called a **Lie algebra**.

Theorem 20. *The ten generators of the Galilean transformation satisfy the commutation relations:*

$$\begin{array}{lll} [J_i, J_j] = i\hbar\,\epsilon_{ijk}J_k\,, & [K_i, K_j] = 0\,, & [J_i, H] = 0\,, \\ [J_i, K_j] = i\hbar\,\epsilon_{ijk}K_k\,, & [P_i, P_j] = 0\,, & [P_i, H] = 0\,, \\ [J_i, P_j] = i\hbar\,\epsilon_{ijk}P_k\,, & [K_i, P_j] = i\hbar\,M\delta_{ij}\,, & [K_i, H] = i\hbar\,P_i\,. \end{array} \tag{9.72}$$

Proof. The proof starts by taking each of the transformations $U(G)$ in theorem 19 to be infinitesimal. These infinitesimal transformations have *nothing* to do with the infinitesimal transformations in the previous theorem — they are different transformations. We start with Eq. (9.59) where we find, to first order:

$$\begin{aligned} &\left\{1 - \frac{i}{\hbar}\,J_k\Delta\theta_k - \frac{i}{\hbar}\,K_k\Delta v_k + \frac{i}{\hbar}\,P_k\Delta a_k - \frac{i}{\hbar}\,H\Delta\tau + \cdots\right\} \\ &\qquad\times\; J_i\left\{1 + \frac{i}{\hbar}\,J_k\Delta\theta_k + \frac{i}{\hbar}\,K_k\Delta v_k - \frac{i}{\hbar}\,P_k\Delta a_k + \frac{i}{\hbar}\,H\Delta\tau + \cdots\right\} \\ &= J_i + \epsilon_{ijk}J_j\Delta\theta_k + \epsilon_{ijk}K_j\Delta v_k + \epsilon_{kji}\Delta a_k P_j + \cdots\,. \end{aligned}$$

Comparing coefficients of $\Delta\theta_k$, Δv_k, Δa_k, and $\Delta\tau$, we find the commutators of J_i with all the other generators:

$$\begin{array}{ll} [J_i, J_j] = i\hbar\,\epsilon_{ijk}J_k\,, & [J_i, P_j] = i\hbar\,\epsilon_{ijk}P_k\,, \\ [J_i, K_j] = i\hbar\,\epsilon_{ijk}K_k\,, & [J_i, H] = 0\,. \end{array}$$

From (9.60), we find, to first order:

$$\begin{aligned} &\left\{1 - \frac{i}{\hbar}\,J_k\Delta\theta_k - \frac{i}{\hbar}\,K_k\Delta v_k + \frac{i}{\hbar}\,P_k\Delta a_k - \frac{i}{\hbar}\,H\Delta\tau + \cdots\right\} \\ &\qquad\times\; K_i\left\{1 + \frac{i}{\hbar}\,J_k\Delta\theta_k + \frac{i}{\hbar}\,K_k\Delta v_k - \frac{i}{\hbar}\,P_k\Delta a_k + \frac{i}{\hbar}\,H\Delta\tau + \cdots\right\} \\ &= K_i + \epsilon_{ijk}K_j\Delta\theta_k + M\Delta a_i - P_i\Delta\tau + \cdots\,, \end{aligned}$$

from which we find the commutators of K_i will all the generators. In addition to the ones found above, we get:

$$[K_i, K_j] = 0\,, \qquad [K_i, P_j] = i\hbar\,M\delta_{ij}\,, \qquad [K_i, H] = i\hbar\,P_i\,.$$

The commutators of P_i with the generators are found from (9.61). We find, to first order:

$$\begin{aligned} &\left\{1 - \frac{i}{\hbar}\,J_k\Delta\theta_k - \frac{i}{\hbar}\,K_k\Delta v_k + \frac{i}{\hbar}\,P_k\Delta a_k - \frac{i}{\hbar}\,H\Delta\tau + \cdots\right\} \\ &\qquad\times\; P_i\left\{1 + \frac{i}{\hbar}\,J_k\Delta\theta_k + \frac{i}{\hbar}\,K_k\Delta v_k - \frac{i}{\hbar}\,P_k\Delta a_k + \frac{i}{\hbar}\,H\Delta\tau + \cdots\right\} \\ &= P_i + \epsilon_{ijk}P_j\Delta\theta_k + M\Delta v_i + \cdots\,, \end{aligned}$$

from which we find the commutators of K_i with all the generators. In addition to the ones found above, we get:

$$[P_i, P_j] = 0\,, \qquad [P_i, H] = 0\,.$$

The last commmutation relations of H with the generators confirm the previous results. This completes the proof. □

The phase parameter M is called a **central charge** of the Galilean algebra.

9.2.4 Center of mass operator

For $M \neq 0$, it is useful to define operators which describes the location and velocity of the center of mass:

Definition 19. The center of mass operator $\mathbf{X}$ is defined at $t = 0$ by $\mathbf{X} = \mathbf{K}/M$. We also define the velocity of the center of mass as $\mathbf{V} = \mathbf{P}/M$.

If no external forces act on the system, the center of mass changes in time according to:

$$\mathbf{X}(t) = \mathbf{X} + \mathbf{V}\,t\,. \tag{9.73}$$

There can still be *internal* forces acting on various parts of the system: we only assume here that the center of mass of the system as a whole moves force free. Using the transformation rules from Theorem 19, $\mathbf{X}(t)$ transforms according to:

$$\begin{aligned} U^\dagger(G)\,\mathbf{X}(t')\,U(G) &= U^\dagger(G)\,\{\,\mathbf{K} + \mathbf{P}\,(t+\tau)\,\}\,U(G)/M \\ &= R\{\mathbf{K} - (\mathbf{P} + M\bar{\mathbf{v}})\,\tau + M\bar{\mathbf{a}} + \mathbf{P}\,(t+\tau) + M\bar{\mathbf{v}}\,(t+\tau)\}/M \\ &= R\{\mathbf{K} + \mathbf{P}\,t\}/M + \mathbf{v}\,t + \mathbf{a} \\ &= R\mathbf{X}(t) + \mathbf{v}\,t + \mathbf{a}\,, \qquad \text{where } t' = t + \tau. \end{aligned} \tag{9.74}$$

Differentiating (9.74) with respect to t', we find:

$$\begin{aligned} U^\dagger(G)\,\dot{\mathbf{X}}(t')\,U(G) &= R\dot{\mathbf{X}}(t) + \mathbf{v}\,, \\ U^\dagger(G)\,\ddot{\mathbf{X}}(t')\,U(G) &= R\ddot{\mathbf{X}}(t)\,, \end{aligned}$$

so the acceleration of the center of mass is an invariant.

We can rewrite the transformation rules and commutation relations of the generators of the Galilean group using $\mathbf{X} = \mathbf{K}/M$ and $\mathbf{V} = \mathbf{P}/M$ rather than $\mathbf{K}$ and $\mathbf{P}$. From Eqs. (9.59–9.62), we find:

$$\begin{aligned} U^\dagger(G)\,\mathbf{J}\,U(G) &= R\{\mathbf{J} + M\mathbf{X}\times\bar{\mathbf{v}} + M\,\bar{\mathbf{a}}\times(\mathbf{V}+\bar{\mathbf{v}})\} \\ &= R\{\mathbf{J} + M\,(\mathbf{X}+\bar{\mathbf{a}})\times\bar{\mathbf{v}} + M\,\bar{\mathbf{a}}\times\mathbf{V}\}\,, \\ U^\dagger(G)\,\mathbf{X}\,U(G) &= R\{\mathbf{X} - (\mathbf{V}+\bar{\mathbf{v}})\,\tau + \bar{\mathbf{a}}\}\,, \\ U^\dagger(G)\,\mathbf{V}\,U(G) &= R\{\mathbf{V}+\bar{\mathbf{v}}\}\,, \\ U^\dagger(G)\,H\,U(G) &= H + M\bar{\mathbf{v}}\cdot\mathbf{V} + \tfrac{1}{2}Mv^2\,. \end{aligned} \tag{9.75}$$

where $\bar{\mathbf{v}} = R^{-1}(\mathbf{v})$ and $\bar{\mathbf{a}} = R^{-1}(\mathbf{a})$. Eqs. (9.72) become:

$$\begin{aligned} &[J_i, J_j] = i\hbar\,\epsilon_{ijk}J_k\,, && [X_i, X_j] = 0\,, && [J_i, H] = 0\,, \\ &[J_i, X_j] = i\hbar\,\epsilon_{ijk}X_k\,, && [P_i, P_j] = 0\,, && [P_i, H] = 0\,, \\ &[J_i, P_j] = i\hbar\,\epsilon_{ijk}P_k\,, && [X_i, P_j] = i\hbar\,\delta_{ij}\,, && [X_i, H] = i\hbar\,V_i\,. \end{aligned} \tag{9.76}$$

Remark 18. For a single particle, the center of mass operator *is* the operator which describes the location of the particle. The existence of such an operator means that we can localize a particle with a measurement of **X**. The commutation relations between **X** and the other generators are as we might expect from the *canonical quantization* postulates which we study in the next chapter. Here, we have obtained these quantization rules directly from the generators of the Galilean group, and from our point of view, they are consequences of requiring Galilean symmetry for the particle system, and are not *additional* postulates of quantum theory. We shall see in a subsequent chapter how to construct quantum mechanics from classical actions.

Remark 19. Since in the Cartesian system of coordinates, **X** and **P** are Hermitian operators, we can always write an eigenvalue equation for them:

$$\mathbf{X}\,|\,\mathbf{x}\,\rangle = \mathbf{x}\,|\,\mathbf{x}\,\rangle\,, \tag{9.77}$$

$$\mathbf{P}\,|\,\mathbf{p}\,\rangle = \mathbf{p}\,|\,\mathbf{p}\,\rangle\,, \tag{9.78}$$

where x_i and p_i are real continuous numbers in the range $-\infty < x_i < \infty$ and $-\infty < p_i < \infty$. In Section 9.4 below, we will find a relationship between these two different basis sets.

9.2.5 Casimir invariants

Casimir operators are operators that are invariant under the transformation group and commute with all the generators of the group. The Galilean transformation is rank two, so we know from a general theorem in group theory that there are just two Casimir operators. These will turn out to be what we will call the **internal energy** W and the magnitude of the **spin S**, or internal angular momentum. We start with the internal energy operator.

Definition 20 (Internal energy)**.** For $M \neq 0$, we define the internal energy operator W by:

$$W = H - \frac{P^2}{2M}\,. \tag{9.79}$$

Theorem 21. *The internal energy, defined Eq.* (9.79)*, is invariant under Galilean transformations:*

Proof. Using Theorem 19, we have:

$$\begin{aligned} U^\dagger(G)\,W\,U(G) &= H + \bar{\mathbf{v}}\cdot\mathbf{P} + \tfrac{1}{2}Mv^2 - \frac{[R\{\mathbf{P}+M\,\bar{\mathbf{v}}\}]^2}{2M} \\ &= H - \frac{P^2}{2M} = W\,, \end{aligned}$$

as required. □

The internal energy operator W is Hermitian and commutes with all the group generators, its eigenvalues w can be any real number. So we can write:

$$H = W + \frac{P^2}{2M}\,. \tag{9.80}$$

The orbital and spin angular momentum operators are defined by:

Definition 21 (Orbital angular momentum)**.** For $M \neq 0$, we define the orbital angular momentum by:

$$\mathbf{L} = \mathbf{X}\times\mathbf{P} = (\mathbf{K}\times\mathbf{P})/M\,. \tag{9.81}$$

The orbital angular momentum of the system is independent of time:

$$\mathbf{L}(t) = \mathbf{X}(t)\times\mathbf{P}(t) = \{\mathbf{X}+\mathbf{P}t/M\}\times\mathbf{P} = \mathbf{X}\times\mathbf{P} = \mathbf{L}\,. \tag{9.82}$$

Definition 22 (Spin)**.** For $M \neq 0$, we define the spin, or internal angular momentum by:

$$\mathbf{S} = \mathbf{J} - \mathbf{L}\,, \tag{9.83}$$

where $\mathbf{L}$ is defined in Eq. (9.81).

The spin is what is left over after subracting the orbital angular momentum from the total angular momentum. Since the orbital angular momentum is not defined for $M = 0$, the same is true for the spin operator. However for $M \neq 0$, we can write:

$$\mathbf{J} = \mathbf{L} + \mathbf{S}\,. \tag{9.84}$$

The following theorem describes the transformation properties of the orbital and spin operators.

Theorem 22. *The orbital and spin operators transform under Galilean transformations according to the rule:*

$$U^\dagger(G)\,\mathbf{L}\,U(G) = R\{\,\mathbf{L} + \mathbf{X}\times(\,M\bar{\mathbf{v}}\,) + (\,\bar{\mathbf{a}} - (\,\mathbf{V} + \bar{\mathbf{v}}\,)\,\tau\,)\times\mathbf{P}\,\}\,, \tag{9.85}$$

$$U^\dagger(G)\,\mathbf{S}\,U(G) = R\{\,\mathbf{S}\,\}\,, \tag{9.86}$$

and obeys the commutation relations:

$$[\,L_i, L_j\,] = i\hbar\,\epsilon_{ijk}L_k\,, \qquad [\,S_i, S_j\,] = i\hbar\,\epsilon_{ijk}S_k\,, \qquad [\,L_i, S_j\,] = 0\,. \tag{9.87}$$

Proof. The orbital results are easy to prove using results from Eqs. (9.75). For the spin, using theorem 19, we find:

$$\begin{aligned}
U^\dagger(G)\,\mathbf{S}\,U(G) &= R\{\,\mathbf{J} + \mathbf{K}\times\bar{\mathbf{v}} + \bar{\mathbf{a}}\times(\mathbf{P} + M\,\bar{\mathbf{v}})\,\}\\
&\quad - R\{\,\mathbf{K} - (\mathbf{P} + M\,\bar{\mathbf{v}})\,\tau + M\,\bar{\mathbf{a}}\,\}\times R\{\mathbf{P} + M\,\bar{\mathbf{v}}\}/M\\
&= R\{\,\mathbf{J} + \mathbf{K}\times\bar{\mathbf{v}} + \bar{\mathbf{a}}\times(\mathbf{P} + M\,\bar{\mathbf{v}})\\
&\quad - [\,\mathbf{K} - (\mathbf{P} + M\,\bar{\mathbf{v}})\,\tau + M\,\bar{\mathbf{a}}\,]\times[\,\mathbf{P} + M\,\bar{\mathbf{v}}\,]/M\,\}\\
&= R\{\,\mathbf{J} - (\mathbf{K}\times\mathbf{P})/M\,\} = R\{\,\mathbf{S}\,\}\,,
\end{aligned}$$

as required. The commutator $[\,L_i, J_j\,] = 0$ is easy to establish. For $[\,L_i, L_j\,]$, we note that:

$$\begin{aligned}
[\,L_i, L_j\,] &= \epsilon_{inm}\epsilon_{jn'm'}\,[\,X_nP_m, X_{n'}P_{m'}\,]\\
&= \epsilon_{inm}\epsilon_{jn'm'}\,\{\,X_{n'}\,[\,X_n, P_{m'}\,]\,P_m + X_n\,[\,P_m, X_{n'}\,]\,P_{m'}\,\}\\
&= i\hbar\,\epsilon_{inm}\epsilon_{jn'm'}\,\{\,\delta_{n,m'}\,X_{n'}\,P_m - \delta_{n',m}\,X_n\,P_{m'}\,\}\\
&= i\hbar\,\{\,\epsilon_{inm}\epsilon_{jn'n}\,X_{n'}\,P_m - \epsilon_{inm}\epsilon_{jmm'}\,X_n\,P_{m'}\,\}\\
&= i\hbar\,\{\,(\,\delta_{mj}\delta_{in'} - \delta_{mn'}\delta_{ij}\,)\,X_{n'}\,P_m - (\,\delta_{im'}\delta_{nj} - \delta_{ij}\delta_{nm'}\,)\,X_n\,P_{m'}\,\}\\
&= i\hbar\,\{\,X_i\,P_j - \delta_{ij}\,(\,X_m\,P_m\,) - X_j\,P_i + \delta_{ij}\,(\,X_n\,P_n\,)\,\}\\
&= i\hbar\,\{\,X_i\,P_j - X_j\,P_i\,\} = i\hbar\,\epsilon_{ijk}\,L_k\,,
\end{aligned} \tag{9.88}$$

as required. The last commutator $[\,S_i, S_j\,]$ follows directly from the commutator results for J_i and L_i. □

Remark 20. Additionally, we note that $[\,S_i, X_j\,] = [\,S_i, P_j\,] = [\,S_i, H\,] = 0$.

Remark 21. So this theorem showns that even under boosts and translations, in addition to rotations, the spin operator is sensitive only to the rotation of the coordinate system, which is not true for either the orbital angular momentum or the total angular momentum operators. However the square of the spin vector operator S^2, *is* invariant under general Galilean transformations,

$$U^{-1}(G)\,S^2\,U(G) = S^2\,, \tag{9.89}$$

and is the second Casimir invariant. In Section **??**, we will find that the possible eigenvalues of S^2 are given by: $s = 0, 1/2, 1, 3/2, 2, \ldots$.

Remark 22. To summerize this section, we have found two hermitian Casimir operators, W and S^2, which are invariant under the group $\mathcal{G}$. We can therefore label the irreducible representations of $\mathcal{G}$ by the set of quantities: $[M|w,s]$, where w and s label the eigenvalues of these operators, and M the central charge.

So we can find common eigenvectors of W, S^2, and either $\mathbf{X}$ or $\mathbf{P}$. We write these as:

$$|\,[M|w,s];\mathbf{x},\sigma\,\rangle\,, \qquad \text{and} \qquad |\,[M|w,s];\mathbf{p},\sigma\,\rangle\,. \tag{9.90}$$

Here σ labels the component of spin. The latter eigenvector is also an eigenvector of H, with eigenvalue:

$$H\,|\,[M|w,s];\mathbf{p},\sigma\,\rangle = E_{w,p}\,|\,[M|w,s];\mathbf{p},\sigma\,\rangle\,, \qquad E_{w,p} = w + \frac{p^2}{2M}\,. \tag{9.91}$$

We discuss the massless case in Section 9.2.8.

9.2.6 Extension of the Galilean group

If we wish, we may extend the Galilean group by considering M to be a generator of the group. This is because the phase factor $\phi(G',G)$ is linear in M and M commutes with all elements of the group. Thus we can invent a new group element η and write:

$$\tilde{G} = (G,\eta) = (R,\mathbf{v},\mathbf{a},\tau,\eta)\,, \tag{9.92}$$

and which transforms according to the rule:

$$\tilde{G}'\tilde{G} = (G'G,\eta'+\eta+\xi(G',G))\,, \tag{9.93}$$

where $\xi(G',G)$ is the coefficient of M in (9.26)

$$\xi(G',G) = -\frac{1}{2}\{\mathbf{v}'\cdot R'\,\mathbf{v}\,\tau + \mathbf{a}'\cdot R'\,\mathbf{v} - \mathbf{v}'\cdot R'\,\mathbf{a}\}\,. \tag{9.94}$$

The infinitesimal unitary operators in Hilbert space become:

$$\tilde{U}(1+\Delta\tilde{G}) = 1 + \frac{i}{\hbar}\,\{\,\mathbf{J}\cdot\hat{\mathbf{n}}\,\theta + \mathbf{K}\cdot\mathbf{v} - \mathbf{P}\cdot\mathbf{a} + H\tau + M\eta\,\} + \cdots\,, \tag{9.95}$$

and since M is now regarded as a generator and η as a group element, the extended eleven parameter Galilean group can now be represented as a *true* unitary representation rather than a projective representation: the phase factor has been redefined as a transformation property of the extended group element η, and the phase M redefined as a operator.

For the extended Galilean group $\tilde{\mathcal{G}}$ with $M\neq 0$, the largest abelian invariant subgroup is now the five dimensional subgroup $\tilde{\mathcal{C}} = [\mathbf{P},H,M]$ generating space and time translations plus η. The abelian invariant subgroup of the factor group $\tilde{\mathcal{G}}/\tilde{\mathcal{C}}$ is then the three parameter subgroup $\mathcal{V} = [\mathbf{K}]$ generating boosts, leaving the semi-simple three-dimensional group of rotations $\mathcal{R} = [R]$. So the extended Galilean group has the product structure:

$$\tilde{\mathcal{G}} = (\mathcal{R}\times\mathcal{V})\times\tilde{\mathcal{C}}\,. \tag{9.96}$$

Here the subgroup $\mathcal{R}\times\mathcal{V} = [\mathbf{J},\mathbf{K}]$ generates the six dimensional group of rotations and boosts.

9.2.7 Finite dimensional representations

We examine in this section finite dimensional representations of the subgroup $\mathcal{R}\times\mathcal{V} = [\mathbf{J},\mathbf{K}]$ of rotations and boosts. These generators obey the subalgebra:

$$[\,J_i,J_j\,] = i\hbar\,\epsilon_{ijk}J_k\,, \qquad [\,J_i,K_j\,] = i\hbar\,\epsilon_{ijk}K_k\,, \qquad [\,K_i,K_j\,] = 0\,. \tag{9.97}$$

In order to emphasize that what we are doing here is completely *classical*, let us define:

$$J_i = \frac{\hbar}{2}\,\Sigma_i\,, \qquad K_i = \frac{\hbar}{2}\,\Gamma_i\,, \tag{9.98}$$

in which case Σ_i and Γ_i satisfy the algebra:

$$[\,\Sigma_i, \Sigma_j\,] = 2\,i\,\epsilon_{ijk}\Sigma_k\,, \qquad [\,\Sigma_i, \Gamma_j\,] = 2\,i\,\epsilon_{ijk}\Gamma_k\,, \qquad [\,\Gamma_i, \Gamma_j\,] = 0\,. \tag{9.99}$$

which eliminates $\hbar$. It is simple to find a 4×4 matrix representation of Σ_i and Γ_i. We find two such complimentary representations:

$$\Sigma_i = \begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix}\,, \qquad \Gamma_i^{(+)} = \begin{pmatrix}0 & 0\\ \sigma_i & 0\end{pmatrix}\,, \qquad \Gamma_i^{(-)} = \begin{pmatrix}0 & \sigma_i\\ 0 & 0\end{pmatrix}\,, \tag{9.100}$$

both of which satisfy the set (9.99):

$$[\,\Sigma_i, \Gamma_j^{(\pm)}\,] = 2\,i\,\epsilon_{ijk}\Gamma_k^{(\pm)}\,, \qquad \Gamma_i^{(\pm)}\Gamma_j^{(\pm)} = 0\,. \tag{9.101}$$

We also find:

$$[\,\Gamma_i^{(+)}, \Gamma_j^{(-)}\,] = \delta_{ij}\,I + i\,\epsilon_{ijk}\,\Sigma_k\,. \tag{9.102}$$

In addition, $[\,\Gamma_i^{(-)}\,]^\dagger = \Gamma_i^{(+)}$ so $\Gamma_i^{(\pm)}$ is *not* Hermitian. Nevertheless, we can define finite transformations by exponentiation. Let us define a rotation operator $U(R)$ by:

$$U(R) = e^{i\hat{\mathbf{n}}\cdot\boldsymbol{\Sigma}\,\theta/2} = I\,\cos\theta/2 + i(\hat{\mathbf{n}}\cdot\boldsymbol{\Sigma})\,\sin\theta/2\,, \tag{9.103}$$

and boost operators $V^{(\pm)}(\mathbf{v})$ by:

$$V^{(+)}(\mathbf{v}) = e^{\mathbf{v}\cdot\boldsymbol{\Gamma}^{(+)}/2} = I + \mathbf{v}\cdot\boldsymbol{\Gamma}^{(+)}/2 = \begin{pmatrix}1 & 0\\ \boldsymbol{\sigma}\cdot\mathbf{v}/2 & 1\end{pmatrix}\,, \tag{9.104}$$

and

$$V^{(-)}(\mathbf{v}) = e^{\mathbf{v}\cdot\Gamma^{(-)}/2} = I + \mathbf{v}\cdot\boldsymbol{\Gamma}^{(-)}/2 = \begin{pmatrix}1 & \boldsymbol{\sigma}\cdot\mathbf{v}/2\\ 0 & 1\end{pmatrix}\,. \tag{9.105}$$

These last two equations follow from the fact that $\Gamma_i^{(\pm)}\Gamma_j^{(\pm)} = 0$. For this same reason,

$$V^{(\pm)}(\mathbf{v}')\,V^{(\pm)}(\mathbf{v}) = V^{(\pm)}(\mathbf{v}'+\mathbf{v})\,. \tag{9.106}$$

We can easily construct the inverses of $V^{(\pm)}(\mathbf{v})$. We find:

$$[\,V^{(\pm)}(\mathbf{v})\,]^{-1} = V^{(\pm)}(-\mathbf{v}) = e^{-\mathbf{v}\cdot\Gamma^{(\pm)}/2} = I - \mathbf{v}\cdot\boldsymbol{\Gamma}^{(\pm)}\,. \tag{9.107}$$

So the inverses of $V^{(\pm)}(\mathbf{v})$ are *not* the adjoints. This means that the $V^{(\pm)}(\mathbf{v})$ operators are *not* unitary.

We now define combined rotation and boost operators by:

$$\Lambda^{(\pm)}(R,\mathbf{v}) = V^{(\pm)}(\mathbf{v})\,U(R)\,, \qquad [\,\Lambda^{(\pm)}(R,\mathbf{v})\,]^{-1} = U^\dagger(R)\,[\,V^{(\pm)}(\mathbf{v})\,]^{-1} = U^\dagger(R)\,V^{(\mp)}(\mathbf{v})\,. \tag{9.108}$$

We find the results:

$$\begin{aligned}
U^\dagger(R)\,\Sigma_i\,U(R) &= R_{ij}\,\Sigma_j\,,\\
U^\dagger(R)\,\Gamma_i^{(\pm)}\,U(R) &= R_{ij}\,\Gamma_j^{(\pm)}\,,\\
[\,V^{(\pm)}(\mathbf{v})\,]^{-1}\,\Sigma_i\,V^{(\pm)}(\mathbf{v}) &= \Sigma_i - 2i\,\epsilon_{ijk}\,\Gamma_j^{(\pm)}\,v_k\,,\\
[\,V^{(\pm)}(\mathbf{v})\,]^{-1}\,\Gamma_i^{(\pm)}\,V^{(\pm)}(\mathbf{v}) &= \Gamma_i^{(\pm)}\,,\\
U^\dagger(R)\,V^{(\pm)}(\mathbf{v})\,U(R) &= V^{(\pm)}(R^{-1}(\mathbf{v}))\,.
\end{aligned} \tag{9.109}$$

So for the combined transformation,

$$\begin{aligned}[\,\Lambda^{(\pm)}(R,\mathbf{v})\,]^{-1}\,\boldsymbol{\Sigma}\,\Lambda^{(\pm)}(R,\mathbf{v}) &= R(\,\boldsymbol{\Sigma}\,) - 2i\,R(\,\boldsymbol{\Gamma}^{(\pm)}\,)\times\mathbf{v}\,,\\ [\,\Lambda^{(\pm)}(R,\mathbf{v})\,]^{-1}\,\boldsymbol{\Gamma}^{(\pm)}\,\Lambda^{(\pm)}(R,\mathbf{v}) &= R(\,\boldsymbol{\Gamma}^{(\pm)}\,)\,.\end{aligned} \tag{9.110}$$

Comparing (9.110) with the transformations of $\mathbf{J}$ and $\mathbf{K}$ in Theorem 19, we see that $\Lambda^{(\pm)}(R,\mathbf{v})$ are adjoint representations of the subgroup rotations and boosts, although not unitary ones. The replacement $\mathbf{v}\to -i\,\mathbf{v}$ is a reflection of the fact that $V^{(\pm)}(\mathbf{v})$ is not unitary. The $\Lambda^{(\pm)}(R,\mathbf{v})$ matrices are faithful representations of the $(R,\mathbf{v})$ subgroup of the Galilean group:

$$\begin{aligned}\Lambda^{(\pm)}(R',\mathbf{v}')\,\Lambda^{(\pm)}(R,\mathbf{v}) &= V^{(\pm)}(\mathbf{v}')\,U(R')\,V^{(\pm)}(\mathbf{v})\,U(R)\\ &= V^{(\pm)}(\mathbf{v}')\,\{\,U(R')\,V^{(\pm)}(\mathbf{v})\,U^{\dagger}(R')\,\}\,U(R')\,U(R)\\ &= V^{(\pm)}(\mathbf{v}')\,V^{(\pm)}(R'(\mathbf{v}))\,U(R'R) = V^{(\pm)}(\mathbf{v}'+R'(\mathbf{v}))\,U(R'R)\\ &= \Lambda^{(\pm)}(R'R,\mathbf{v}'+R'(\mathbf{v}))\,.\end{aligned} \tag{9.111}$$

We can, in fact, display an explicit Galilean transformation for the subgroup consisting of the $(R,\mathbf{v})$ elements. Let us define two 4×4 matrices $X^{(\pm)}(\mathbf{x},t)$ by:

Definition 23.

$$X^{(+)}(\mathbf{x},t) = \begin{pmatrix} t & 0\\ \mathbf{x}\cdot\boldsymbol{\sigma} & -t\end{pmatrix}\,,\qquad X^{(-)}(\mathbf{x},t) = \begin{pmatrix} -t & \mathbf{x}\cdot\boldsymbol{\sigma}\\ 0 & t\end{pmatrix}\,. \tag{9.112}$$

Then we can prove the following theorem:

Theorem 23. *The matrices $X^{(\pm)}(\mathbf{x},t)$ transform under the subgroup of rotations and boosts according to:*

$$\Lambda^{(\pm)}(R,\mathbf{v})\,X^{(\pm)}(\mathbf{x},t)\,[\,\Lambda^{(\pm)}(R,\mathbf{v})\,]^{-1} = X^{(\pm)}(\mathbf{x}',t')\,, \tag{9.113}$$

where $\mathbf{x}' = R(\mathbf{x}) + \mathbf{v}t$ and $t' = t$.

Proof. This remarkable result is an alternative way of writing Galilean transformations for the subgroup of rotations and boosts in terms of transformations of 4×4 matrices in the "adjoint" representation. With the above definitions, the proof is straightforward and is left for the reader. □

Exercise 9. Prove Theorem 23.

In this section, we have found two 4×4 dimensional matrix representations of the Galilean group. These representations turned out not to be unitary. Finite dimensional representations of the Lorentz group in relativistic theories are also not unitary. Nevertheless, finite representations of the Galilean group will be useful when discussing wave equations.

9.2.8 The massless case

When $M = 0$, the phase for unitary representations of the Galilean group vanish, and the representation becomes a faithful one, which is simpler. For this case, the generators transform according to the equations:

$$\begin{aligned}U^{\dagger}(G)\,\mathbf{J}\,U(G) &= R\{\mathbf{J}+\mathbf{K}\times\bar{\mathbf{v}}+\bar{\mathbf{a}}\times\mathbf{P}\}\,,\\ U^{\dagger}(G)\,\mathbf{K}\,U(G) &= R\{\mathbf{K}-\mathbf{P}\,\tau\}\,,\\ U^{\dagger}(G)\,\mathbf{P}\,U(G) &= R\{\mathbf{P}\}\,,\\ U^{\dagger}(G)\,H\,U(G) &= H+\bar{\mathbf{v}}\cdot\mathbf{P}\,.\end{aligned} \tag{9.114}$$

where $\bar{\mathbf{v}} = R^{-1}(\mathbf{v})$ and $\bar{\mathbf{a}} = R^{-1}(\mathbf{a})$. The generators obey the algebra:

$$\begin{aligned} [J_i, J_j] &= i\hbar\,\epsilon_{ijk} J_k\,, & [K_i, K_j] &= 0\,, & [J_i, H] &= 0\,, \\ [J_i, K_j] &= i\hbar\,\epsilon_{ijk} K_k\,, & [P_i, P_j] &= 0\,, & [P_i, H] &= 0\,, \\ [J_i, P_j] &= i\hbar\,\epsilon_{ijk} P_k\,, & [K_i, P_j] &= 0\,, & [K_i, H] &= i\hbar\,P_i\,. \end{aligned} \tag{9.115}$$

We first note that $\mathbf{P}$ simply rotates like a vector under the full group, so P^2 is the first Casimir invariant.

We also note that if we define $\mathbf{W} = \mathbf{K} \times \mathbf{P}$, then

$$U^\dagger(G)\,\mathbf{W}\,U(G) = R\{\mathbf{K} - \mathbf{P}\,\tau\} \times R\{\mathbf{P}\} = R\{\mathbf{K} \times \mathbf{P}\} = R\{\,\mathbf{W}\}\,. \tag{9.116}$$

So $\mathbf{W}$ is a second vector which simply rotates like a vector under the full group, so W^2 is also an invariant. We also note that $\mathbf{W}$ is perpendicular to both $\mathbf{P}$ and $\mathbf{K}$: $\mathbf{W}\cdot\mathbf{P} = \mathbf{W}\cdot\mathbf{K} = 0$. Note that $\mathbf{W}$ does not satisfy angular momentum commutator relations.

9.3 Time translations

We have only constructed the unitary operator $U(1 + \Delta G)$ for infinitesimal Galilean transformations. Since the generators do not commute, we cannot construct the unitary operator $U(G)$ for a *finite* Galilean transformation by application of a series of infinitesimal ones. However we can easily construct the unitary operator $U(G)$ for restricted Galilean transformations, like time, space, and boost transformations alone. We do this in the next two sections.

The unitary operator for pure time translations is given by:

$$U_{_H}(\tau) = \lim_{N\to\infty}\left[1 + \frac{i}{\hbar}\frac{H\,\tau}{N}\right]^N = e^{iH\,\tau/\hbar}\,. \tag{9.117}$$

It time-translates the operator $\mathbf{X}(t)$ by an amount τ:

$$U^\dagger_{_H}(\tau)\,\mathbf{X}(t')\,U_{_H}(\tau) = \mathbf{X}(t)\,, \qquad t' = t + \tau\,, \tag{9.118}$$

and leaves P unchanged:

$$U^\dagger_{_H}(\tau)\,\mathbf{P}\,U_{_H}(\tau) = \mathbf{P}\,. \tag{9.119}$$

Invariance of the laws of nature under time translation is a statement of the fact that an experiment with particles done today will give the same results as an experiment done yesterday — there is no way of measuring absolute time.

We first consider transformations to a frame where we have set the clocks to zero. That is, we put $t' = 0$ so that $\tau = -t$. Then (9.118) becomes:

$$\mathbf{X}(t) = U_{_H}(t)\,\mathbf{X}\,U^\dagger_{_H}(t) = \mathbf{X} + \mathbf{V}\,t\,. \tag{9.120}$$

where $\mathbf{X} = \mathbf{K}/M$ and $\mathbf{V} = \mathbf{P}/M$. From Eq. (9.120), we find:

$$\mathbf{X}(t)\,\{U_{_H}(t)\,|\,\mathbf{x}\,\rangle\} = U_{_H}(t)\,\mathbf{X}\,|\,\mathbf{x}\,\rangle = \mathbf{x}\,\{U_{_H}(t)\,|\,\mathbf{x}\,\rangle\}\,. \tag{9.121}$$

So if we define the ket $|\,\mathbf{x}, t\,\rangle$ by:

$$|\,\mathbf{x}, t\,\rangle = U_{_H}(t)\,|\,\mathbf{x}\,\rangle = e^{iHt/\hbar}\,|\,\mathbf{x}\,\rangle\,, \tag{9.122}$$

then (9.121) becomes an eigenvalue equation for the operator $\mathbf{X}(t)$ at time t:

$$\mathbf{X}(t)\,|\,\mathbf{x}, t\,\rangle = \mathbf{x}\,|\,\mathbf{x}, t\,\rangle\,, \qquad \mathbf{x} \in \mathbb{R}^3\,. \tag{9.123}$$

Note that the *eigenvalue* $\mathbf{x}$ of this equation is not a function of t. It is just a real vector.

From Eq. (9.122), we see that the base vector $|\,\mathbf{x},t\,\rangle$ satisfies a first order differential equation:

$$-i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\,\mathbf{x},t\,\rangle = H\,|\,\mathbf{x},t\,\rangle\,, \tag{9.124}$$

and from (9.120), we obtain Heisenberg's differential equation of motion for $\mathbf{X}(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{X}(t) = [\mathbf{X}(t),H]/i\hbar = \mathbf{P}/M\,. \tag{9.125}$$

The general transformation of the base vectors $|\,\mathbf{x},t\,\rangle$ between two frames, which differ by clock time τ *only*, is given by:

$$|\,\mathbf{x},t'\,\rangle = U_H(t')\,|\,\mathbf{x}\,\rangle = U_H(t')\,U_H^\dagger(t)\,|\,\mathbf{x},t\,\rangle = U_H(\tau)\,|\,\mathbf{x},t\,\rangle\,, \tag{9.126}$$

where $\tau = t'-t$.

The inner product of $|\,\mathbf{x},t\,\rangle$ with an arbitrary vector $|\,\psi\,\rangle$ is given by:

$$\psi(\mathbf{x},t) = \langle\,\mathbf{x},t\,|\,\psi\,\rangle = \langle\,\mathbf{x}\,|U_H^\dagger(t)|\,\psi\,\rangle = \langle\,\mathbf{x}\,|\,\psi(t)\,\rangle\,, \tag{9.127}$$

where the time-dependent "state vector" $|\,\psi(t)\,\rangle$ is defined by:

$$|\,\psi(t)\,\rangle = U_H^\dagger(t)\,|\,\psi\,\rangle = e^{-iHt/\hbar}\,|\,\psi\,\rangle\,. \tag{9.128}$$

This state vector satisfies a differential equation given by:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\,\psi(t)\,\rangle = H\,|\,\psi(t)\,\rangle\,, \tag{9.129}$$

which is called **Schrödinger's** equation. This equation gives the trajectory of the state vector in Hilbert space. Thus, we can consider two pictures: base vectors moving (the Heisenberg picture) or state vector moving (the Schrödinger picture). They are different views of the same physics. From our point of view, *and remarkably*, Schrödinger's equation is a result of requiring Galilean symmetry, and is not a fundamental postulate of the theory.

The state vector in the primed frame is related to that in the unprimed frame by:

$$|\,\psi(t')\,\rangle = U_H^\dagger(t')\,|\,\psi\,\rangle = U_H^\dagger(t')U_H(t)\,|\,\psi(t)\,\rangle = U_H^\dagger(\tau)\,|\,\psi(t)\,\rangle\,, \tag{9.130}$$

We next turn to space translations and boosts.

9.4 Space translations and boosts

The unitary operators for pure space translations and pure boosts are built up of infinitesimal transformations along any path:

$$U_{\mathbf{P}}(\mathbf{a}) = \lim_{N\to\infty}\left[1-\frac{i}{\hbar}\frac{\mathbf{P}\cdot\mathbf{a}}{N}\right]^N = e^{-i\mathbf{P}\cdot\mathbf{a}/\hbar}\,, \tag{9.131}$$

$$U_{\mathbf{X}}(\mathbf{v}) = \lim_{N\to\infty}\left[1+\frac{i}{\hbar}\frac{\mathbf{K}\cdot\mathbf{v}}{N}\right]^N = e^{i\mathbf{K}\cdot\mathbf{v}/\hbar} = e^{iM\mathbf{v}\cdot\mathbf{X}/\hbar}\,, \tag{9.132}$$

The space translation operator $U_{\mathbf{P}}(\mathbf{a})$ is diagonal in momentum eigenvectors, and the boost operator $U_{\mathbf{X}}(\mathbf{v})$ is diagonal in position eigenvectors. From the transformation rules, we have:

$$U_{\mathbf{P}}^\dagger(\mathbf{a})\,\mathbf{X}\,U_{\mathbf{P}}(\mathbf{a}) = \mathbf{X}+\mathbf{a}\,, \tag{9.133}$$

$$U_{\mathbf{X}}^\dagger(\mathbf{v})\,\mathbf{P}\,U_{\mathbf{X}}(\mathbf{v}) = \mathbf{P}+M\mathbf{v}\,. \tag{9.134}$$

Thus $U_{\mathbf{P}}(\mathbf{a})$ translates the position operator and $U_{\mathbf{X}}(\mathbf{v})$ translates the momentum operator. For the eigenvectors, this means that, for the case of no degeneracies,

$$|\,\mathbf{x}'\,\rangle = |\,\mathbf{x}+\mathbf{a}\,\rangle \quad = U_{\mathbf{P}}(\mathbf{a})\,|\,\mathbf{x}\,\rangle\,, \tag{9.135}$$

$$|\,\mathbf{p}'\,\rangle = |\,\mathbf{p}+M\mathbf{v}\,\rangle = U_{\mathbf{X}}(\mathbf{v})\,|\,\mathbf{p}\,\rangle\,, \tag{9.136}$$

In this section, we omit the explicit reference to w. We can find any ket from "standard" kets $|\,\mathbf{x}_0\,\rangle$ and $|\,\mathbf{p}_0\,\rangle$ by translation and boost operators, as we did for time translations. Thus in Eq. (9.135), we set $\mathbf{x} = \mathbf{x}_0 \equiv 0$, and then put $\mathbf{a} \to \mathbf{x}$, and in Eq. (9.136), we set $\mathbf{p} = \mathbf{p}_0 \equiv 0$, and put $\mathbf{v} \to \mathbf{p}/M$. This gives the relations:

$$|\,\mathbf{x}\,\rangle = U_{\mathbf{P}}(\mathbf{x})\,|\,\mathbf{x}_0\,\rangle\,, \tag{9.137}$$

$$|\,\mathbf{p}\,\rangle = U_{\mathbf{X}}(\mathbf{p}/M)\,|\,\mathbf{p}_0\,\rangle\,. \tag{9.138}$$

We can use (9.137) or (9.138) to find a relation between the $|\,\mathbf{x}\,\rangle$ and $|\,\mathbf{p}\,\rangle$ representations. We have:

$$\langle\,\mathbf{x}\,|\,\mathbf{p}\,\rangle = \langle\,\mathbf{x}\,|\,U_{\mathbf{X}}(\mathbf{p}/M)\,|\,\mathbf{p}_0\,\rangle = \langle\,\mathbf{x}_0\,|\,U^{\dagger}_{\mathbf{P}}(\mathbf{x})\,|\,\mathbf{p}\,\rangle = N\,e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}\,,$$

where $N = \langle\,\mathbf{x}_0\,|\,\mathbf{p}\,\rangle = \langle\,\mathbf{x}\,|\,\mathbf{p}_0\,\rangle$.

In this book, we normalize these states according to the rule:

$$\sum_{\mathbf{x}} \to \int \mathrm{d}^3x\,, \tag{9.139}$$

$$\sum_{\mathbf{p}} \to \int \frac{\mathrm{d}^3p}{(2\pi\hbar)^3}\,, \tag{9.140}$$

Then we have the normalizations:

$$\langle\,\mathbf{x}\,|\,\mathbf{x}'\,\rangle = \sum_{\mathbf{p}}\langle\,\mathbf{x}\,|\,\mathbf{p}\,\rangle\langle\,\mathbf{p}\,|\,\mathbf{x}'\,\rangle = \delta(\mathbf{x}-\mathbf{x}')\,, \tag{9.141}$$

$$\langle\,\mathbf{p}\,|\,\mathbf{p}'\,\rangle = \sum_{\mathbf{x}}\langle\,\mathbf{p}\,|\,\mathbf{x}\,\rangle\langle\,\mathbf{x}\,|\,\mathbf{p}'\,\rangle = (2\pi\hbar)^3\,\delta(\mathbf{p}-\mathbf{p}')\,. \tag{9.142}$$

This means that we should take the normalization $N = 1$, so that the Fourier transform pair is given by:

$$\psi(\mathbf{x}) = \langle\,\mathbf{x}\,|\,\psi\,\rangle = \sum_{\mathbf{p}}\langle\,\mathbf{x}\,|\,\mathbf{p}\,\rangle\langle\,\mathbf{p}\,|\,\psi\,\rangle = \int \frac{\mathrm{d}^3p}{(2\pi\hbar)^3}\,e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}\,\tilde{\psi}(\mathbf{p})\,, \tag{9.143}$$

$$\tilde{\psi}(\mathbf{p}) = \langle\,\mathbf{p}\,|\,\psi\,\rangle = \sum_{\mathbf{x}}\langle\,\mathbf{p}\,|\,\mathbf{x}\,\rangle\langle\,\mathbf{x}\,|\,\psi\,\rangle = \int \mathrm{d}^3x\,e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar}\,\psi(\mathbf{x}) \tag{9.144}$$

For pure space translations, $\mathbf{x}' = \mathbf{x}+\mathbf{a}$, wave functions in coordinate space transform according to the rule:

$$\psi'(\mathbf{x}') = \langle\,\mathbf{x}'\,|\,\psi'\,\rangle = \langle\,\mathbf{x}\,|\,U^{\dagger}_{\mathbf{P}}(\mathbf{a})U_{\mathbf{P}}(\mathbf{a})\,|\,\psi\,\rangle = \langle\,\mathbf{x}\,|\,\psi\,\rangle = \psi(\mathbf{x})\,. \tag{9.145}$$

For infinitesimal displacements, $\mathbf{x}' = \mathbf{x}+\Delta\mathbf{a}$, we have, using Taylor's expansion,

$$\begin{aligned}\psi(\mathbf{x}+\Delta\mathbf{a}) = \langle\,\mathbf{x}\,|\,U^{\dagger}_{\mathbf{P}}(\Delta\mathbf{a})\,|\,\psi\,\rangle &= 1 + \frac{i}{\hbar}\,\Delta\mathbf{a}\cdot\langle\,x\,|\mathbf{P}|\,\psi\,\rangle + \cdots \\ &= \left\{1 + \Delta\mathbf{a}\cdot\boldsymbol{\nabla}_{\mathbf{x}} + \cdots\right\}\psi(\mathbf{x})\,.\end{aligned}$$

So the coordinate representation of the momentum operator is:

$$\langle\,\mathbf{x}\,|\,\mathbf{P}\,|\,\psi\,\rangle = \frac{\hbar}{i}\,\boldsymbol{\nabla}_{\mathbf{x}}\,\psi(\mathbf{x})\,. \tag{9.146}$$

In a similar way, for pure boosts, $\mathbf{p}' = \mathbf{p} + M\mathbf{v}$, wave functions in momentum space transforms according to:

$$\tilde{\psi}'(\mathbf{p}') = \langle\, \mathbf{p}' \,|\, \psi' \,\rangle = \langle\, \mathbf{p} \,|\, U^{\dagger}_{\mathbf{X}}(\mathbf{v}) U_{\mathbf{X}}(\mathbf{v}) \,|\, \psi \,\rangle = \langle\, \mathbf{p} \,|\, \psi \,\rangle = \tilde{\psi}(\mathbf{p})\,, \tag{9.147}$$

and we find:

$$\langle\, p \,|\, \mathbf{X} \,|\, \psi \,\rangle = -\frac{\hbar}{i}\, \boldsymbol{\nabla}_{\mathbf{p}}\, \tilde{\psi}(\mathbf{x})\,. \tag{9.148}$$

For the combined unitary operator for space translations and boosts, we note that the combined transformations give: $(1, \mathbf{v}, 0, 0)(1, 0, \mathbf{a}, 0) = (1, \mathbf{v}, \mathbf{a}, 0)$. So, using Bargmann's theorem, Eq. (9.26), for the phase, and Eq. (B.16) in Appendix **??**, we find the results:

$$\begin{aligned} U_{\mathbf{X},\mathbf{P}}(\mathbf{v}, \mathbf{a}) = e^{i(M\mathbf{v}\cdot\mathbf{X} - \mathbf{P}\cdot\mathbf{a})/\hbar} &= e^{+i\frac{1}{2}M\mathbf{v}\cdot\mathbf{a}/\hbar}\, U_{\mathbf{P}}(\mathbf{a})\, U_{\mathbf{X}}(\mathbf{v})\,, \\ &= e^{-i\frac{1}{2}M\mathbf{v}\cdot\mathbf{a}/\hbar}\, U_{\mathbf{X}}(\mathbf{v})\, U_{\mathbf{P}}(\mathbf{a})\,, \end{aligned} \tag{9.149}$$

So for *combined* space translations and boosts we find:

$$\begin{aligned} U_{\mathbf{X},\mathbf{P}}(\mathbf{v}, \mathbf{a}) \,|\, \mathbf{x} \,\rangle &= e^{+i\frac{1}{2}M\mathbf{v}\cdot\mathbf{a}/\hbar}\, U_{\mathbf{P}}(\mathbf{a})\, U_{\mathbf{X}}(\mathbf{v}) \,|\, \mathbf{x} \,\rangle \\ &= e^{+i(M\mathbf{v}\cdot\mathbf{x} + \frac{1}{2}M\mathbf{v}\cdot\mathbf{a})/\hbar}\, U_{\mathbf{P}}(\mathbf{a}) \,|\, \mathbf{x} \,\rangle \\ &= e^{+i(M\mathbf{v}\cdot\mathbf{x} + \frac{1}{2}M\mathbf{v}\cdot\mathbf{a})/\hbar} \,|\, \mathbf{x} + \mathbf{a} \,\rangle \\ U_{\mathbf{X},\mathbf{P}}(\mathbf{v}, \mathbf{a}) \,|\, \mathbf{p} \,\rangle &= e^{-i\frac{1}{2}M\mathbf{v}\cdot\mathbf{a}/\hbar}\, U_{\mathbf{X}}(\mathbf{v})\, U_{\mathbf{P}}(\mathbf{a}) \,|\, \mathbf{p} \,\rangle \\ &= e^{-i(\mathbf{p}\cdot\mathbf{a} + \frac{1}{2}M\mathbf{v}\cdot\mathbf{a})/\hbar}\, U_{\mathbf{X}}(\mathbf{v}) \,|\, \mathbf{p} \,\rangle \\ &= e^{-i(\mathbf{p}\cdot\mathbf{a} + \frac{1}{2}M\mathbf{v}\cdot\mathbf{a})/\hbar} \,|\, \mathbf{p} + M\mathbf{v} \,\rangle\,. \end{aligned}$$

Writing $\mathbf{x}' = \mathbf{x} + \mathbf{a}$ and $\mathbf{p}' = \mathbf{p} + M\mathbf{v}$, and inverting these expressions, we find

$$|\, \mathbf{x}' \,\rangle = e^{-i(M\mathbf{v}\cdot\mathbf{x} + \frac{1}{2}M\mathbf{v}\cdot\mathbf{a})/\hbar}\, U_{\mathbf{X},\mathbf{P}}(\mathbf{v}, \mathbf{a}) \,|\, \mathbf{x} \,\rangle\,, \tag{9.150}$$

$$|\, \mathbf{p}' \,\rangle = e^{+i(\mathbf{p}\cdot\mathbf{a} + \frac{1}{2}M\mathbf{v}\cdot\mathbf{a})/\hbar}\, U_{\mathbf{X},\mathbf{P}}(\mathbf{v}, \mathbf{a}) \,|\, \mathbf{p} \,\rangle\,. \tag{9.151}$$

For combined transformations, wave functions in coordinate and momentum space transform according to the rule:

$$\psi'(\mathbf{x}') = \langle\, \mathbf{x}' \,|\, \psi' \,\rangle = \langle\, \mathbf{x}' \,| U_{\mathbf{X},\mathbf{P}}(\mathbf{v}, \mathbf{a}) \,|\, \psi \,\rangle = e^{+i(M\mathbf{v}\cdot\mathbf{x} + \frac{1}{2}M\mathbf{v}\cdot\mathbf{a})/\hbar}\, \psi(\mathbf{x})\,, \tag{9.152}$$

$$\tilde{\psi}'(\mathbf{p}') = \langle\, \mathbf{p}' \,|\, \psi' \,\rangle = \langle\, \mathbf{p}' \,| U_{\mathbf{X},\mathbf{P}}(\mathbf{v}, \mathbf{a}) \,|\, \psi \,\rangle = e^{-i(\mathbf{p}\cdot\mathbf{a} + \frac{1}{2}M\mathbf{v}\cdot\mathbf{a})/\hbar}\, \tilde{\psi}(\mathbf{p})\,. \tag{9.153}$$

These functions transform like scalars, but with an essential coordinate or momenutm dependent phase, characteristic of Gailiean transformations.

Example 29. It is easy to show that Eq. (9.152), *is* the Fourier transform of (9.153),

$$\begin{aligned} \psi'(\mathbf{x}') &= \int \frac{\mathrm{d}^3 p'}{(2\pi\hbar)^3}\, e^{i\mathbf{p}'\cdot\mathbf{x}'/\hbar}\, \tilde{\psi}'(\mathbf{p}') \\ &= e^{+i(M\mathbf{v}\cdot\mathbf{x} + \frac{1}{2}M\mathbf{v}\cdot\mathbf{a})/\hbar} \int \frac{\mathrm{d}^3 p}{(2\pi\hbar)^3}\, e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}\, \tilde{\psi}(\mathbf{p}) = e^{+i(M\mathbf{v}\cdot\mathbf{x} + \frac{1}{2}M\mathbf{v}\cdot\mathbf{a})/\hbar}\, \psi(\mathbf{x})\,. \end{aligned}$$

as required by Eq. (9.143).

We discuss the case of combined space and time translations with boosts, but without rotations, in Appendix **??**. We turn next to rotations.

9.5 Rotations

In this section, we discuss pure rotations. Because of the importance of rotations and angular momentum in quantum mechanics, this topic is discussed in great detail in Chapter **??**. We will therefore restrict our discussion here to general properties of pure rotations and angular momentum algebra.

9.5.1 The rotation operator

The total angular momentum is the sum of orbital plus spin: $\mathbf{J} = \mathbf{L} + \mathbf{S}$, with $[\, L_i, S_j \,] = 0$. Common eigenvectors of these two operators are then the direct product of these two states:

$$|\, \ell, m_\ell; s, m_s \,\rangle = |\, \ell, m_\ell \,\rangle \, |\, s, m_s \,\rangle \, . \tag{9.154}$$

The rotation operator is given by the combined rotation of orbital and spin operators:

$$U_{\mathbf{J}}(R) = e^{i\hat{\mathbf{n}}\cdot\mathbf{J}\,\theta/\hbar} = e^{i\hat{\mathbf{n}}\cdot\mathbf{L}\,\theta/\hbar}\, e^{i\hat{\mathbf{n}}\cdot\mathbf{S}\,\theta/\hbar} = U_{\mathbf{L}}(R)\, U_{\mathbf{S}}(R) \, . \tag{9.155}$$

The orbital rotation operator acts only on eigenstates of the position operator $\mathbf{X}$, or momentum operator $\mathbf{P}$,

For pure rotations, the rotation operator can be found by N sequential infinitesimal transformations $\Delta\theta = \theta/N$ about a fixed axis $\hat{\mathbf{n}}$:

$$U_{\mathbf{J}}(\hat{\mathbf{n}}, \theta) = \lim_{N\to\infty} \left[1 + \frac{i}{\hbar} \frac{\mathbf{n}\cdot\mathbf{J}\,\theta}{N} \right]^N = e^{i\hat{\mathbf{n}}\cdot\mathbf{J}\,\theta/\hbar} \, . \tag{9.156}$$

For pure rotations, the Galilean phase factor is zero so that we have:

$$U_{\mathbf{J}}(R')\, U_{\mathbf{J}}(R) = U_{\mathbf{J}}(R'R) \, . \tag{9.157}$$

From Theorem 19 and Eq. (9.59), for pure rotations, we have:

$$U^\dagger_{\mathbf{J}}(\hat{\mathbf{n}}, \theta)\, J_i\, U_{\mathbf{J}}(\hat{\mathbf{n}}, \theta) = R_{ij}(\hat{\mathbf{n}}, \theta)\, J_j \equiv J_i(\hat{\mathbf{n}}, \theta) \, . \tag{9.158}$$

We discuss parameterizations of the rotation matrices $R(\hat{\mathbf{n}}, \theta)$ in Appendix **??**. Here $J_i(\hat{\mathbf{n}}, \theta)$ is the i^{th} component of the operator $\mathbf{J}$ evaluated in the rotated system. Setting $i = z$, we find for the z-component:

$$J_z(\hat{\mathbf{n}}, \theta)\, U^\dagger_{\mathbf{J}}(\hat{\mathbf{n}}, \theta) = U^\dagger_{\mathbf{J}}(\hat{\mathbf{n}}, \theta)\, J_z \tag{9.159}$$

We also know that $J^2 = J_x^2 + J_y^2 + J_z^2$ is an invariant:

$$U^\dagger_{\mathbf{J}}(\hat{\mathbf{n}}, \theta)\, J^2\, U_{\mathbf{J}}(\hat{\mathbf{n}}, \theta) = J^2 \, . \tag{9.160}$$

So from Eq. (9.159), we find that:

$$J_z(\hat{\mathbf{n}}, \theta) \left\{ U^\dagger_{\mathbf{J}}(\hat{\mathbf{n}}, \theta) \, |\, j, m \,\rangle \right\} = \hbar\, m \left\{ U^\dagger_{\mathbf{J}}(\hat{\mathbf{n}}, \theta) \, |\, j, m \,\rangle \right\} , \tag{9.161}$$

from which we conclude that the quantity in brackets is an eigenvector of $J_z(\hat{\mathbf{n}}, \theta)$ with eigenvalue $\hbar m$. That is, we can write:

$$|\, j, m(\hat{\mathbf{n}}, \theta) \,\rangle = U^\dagger_{\mathbf{J}}(\hat{\mathbf{n}}, \theta) \, |\, j, m \,\rangle \, . \tag{9.162}$$

It is also an eigenvector of J^2 with eigenvalue $\hbar^2\, j(j+1)$. It is useful to define a rotation matrix $D^{(j)}_{m',m}(\hat{\mathbf{n}}, \theta)$ by:

$$D^{(j)}_{m,m'}(\hat{\mathbf{n}}, \theta) = \langle\, jm \,|\, U_{\mathbf{J}}(\hat{\mathbf{n}}, \theta) \,|\, jm' \,\rangle \, . \tag{9.163}$$

Matrix elements of the rotation operator are diagonal in j. The rotation matrices have the properties:

$$\sum_{m'=-j}^{j} D^{(j)}_{m,m'}(R)\, D^{(j)\,*}_{m'',m'}(R) = \delta_{m,m''}\,, \tag{9.164}$$

$$D^{(j)\,*}_{m,m'}(R) = D^{(j)}_{m',m}(R^{-1}) = (-)^{m'-m}\, D^{(j)}_{-m,-m'}(R)\,. \tag{9.165}$$

We can express $|\,\hat{\mathbf{n}},\theta;j,m\,\rangle$ in terms of the rotation matrices. We write:

$$|\,j,m(\hat{\mathbf{n}},\theta)\,\rangle = \sum_{m'=-j}^{j} D^{(j)\,*}_{m,m'}(\hat{\mathbf{n}},\theta)\,|\,j,m'\,\rangle\,. \tag{9.166}$$

In the coordinate representation of orbital angular momenta, spherical harmonics are defined by: $Y_{\ell,m}(\Omega) = \langle\,\Omega\,|\,\ell,m\,\rangle$. Using Eq. (9.166), we find:

$$\begin{aligned} Y_{\ell,m}(\Omega') &= \langle\,\Omega'\,|\,\ell,m\,\rangle = \langle\,\Omega\,|\,U^{\dagger}_{\mathbf{J}}(\hat{\mathbf{n}},\theta)\,|\,\ell,m\,\rangle \\ &= \sum_{m'=-\ell}^{\ell} D^{(\ell)\,*}_{m,m'}(\hat{\mathbf{n}},\theta)\,\langle\,\Omega\,|\,\ell,m'\,\rangle = \sum_{m'=-\ell}^{\ell} D^{(\ell)\,*}_{m,m'}(\hat{\mathbf{n}},\theta)\,Y_{\ell,m'}(\Omega)\,, \end{aligned} \tag{9.167}$$

where Ω and Ω' are spherical angles of the same point measured in two different coordinate systems, rotated relative to each other.

9.5.2 Rotations of the basis sets

Now $\mathbf{L}$ and therefore $\mathbf{J}$ does *not* commute with either $\mathbf{X}$ or $\mathbf{P}$. Therefore they cannot have common eigenvectors. However $\mathbf{S}$ *does* commute with with both $\mathbf{X}$ or $\mathbf{P}$. Supressing the dependence on w and M, the common eigenvectors are:

$$|\,\mathbf{x},sm\,\rangle\,, \qquad \text{and} \qquad |\,\mathbf{p},sm\,\rangle\,. \tag{9.168}$$

A general rotation of the ket $|\,\mathbf{x},sm\,\rangle$ can be obtained by first translating to the state where $\mathbf{x}=0$, then rotating, and then translating back to a rotated state $\mathbf{x}'=R(\mathbf{x})$. That is, $(R,0,0,0) = (1,\mathbf{x}',0,0)(R,0,0,0)(1,-\mathbf{x},0,0)$. The trick is that the orbital angular momentum operator $\mathbf{L}$ acting on a state with $\mathbf{x}=\mathbf{0}$ gives zero, so on this state $\mathbf{J}=\mathbf{S}$. The phases all work out to be zero in this case, so we find:

$$\begin{aligned} U_{\mathbf{J}}(R)\,|\,\mathbf{x},sm\,\rangle &= U_{\mathbf{P}}(\mathbf{x}')\,U_{\mathbf{J}}(R)\,U_{\mathbf{P}}(-\mathbf{x})\,|\,\mathbf{x},sm\,\rangle \\ &= U_{\mathbf{P}}(\mathbf{x}')\,U_{\mathbf{J}}(R)\,|\,\mathbf{0},sm\,\rangle \\ &= U_{\mathbf{P}}(\mathbf{x}')\,U_{\mathbf{S}}(R)\,|\,\mathbf{0},sm\,\rangle \\ &= \sum_{m'} U_{\mathbf{P}}(\mathbf{x}')\,|\,\mathbf{0},sm'\,\rangle\,D^{(s)}_{m',m}(R) \\ &= \sum_{m'} |\,\mathbf{x}',sm'\,\rangle\,D^{(s)}_{m',m}(R)\,. \end{aligned} \tag{9.169}$$

Inverting this expression, we find:

$$U^{\dagger}_{\mathbf{J}}(R)\,|\,\mathbf{x}',sm'\,\rangle = \sum_{m} D^{(s)*}_{m',m}(R)\,|\,\mathbf{x},sm\,\rangle\,, \tag{9.170}$$

which gives:

$$\psi'_{sm'}(\mathbf{x}') = \sum_{m} D^{(s)}_{m',m}(R)\,\psi_{sm}(\mathbf{x})\,, \tag{9.171}$$

where $\langle\,\mathbf{x},sm\,|\,\psi\,\rangle = \psi_{sm}(\mathbf{x})$ with $|\,\psi'\,\rangle = U(R)\,|\,\psi\,\rangle$.

9.6 General Galilean transformations

The general Galilean transformation for space and time translations and rotations is given by:

$$\begin{aligned} \mathbf{x}' &= R(\mathbf{x}) + \mathbf{v}t + \mathbf{a}\,, \\ t' &= t + \tau\,. \end{aligned} \tag{9.172}$$

Starting from the state $|\, sm; \mathbf{x}, t\,\rangle$, we generate a full Galilean transformation $G = (R, \mathbf{v}, \mathbf{a}, \tau)$ by first doing a time translation back to $t = 0$, a space translation back to the origin $\mathbf{x} = \mathbf{0}$, then a rotation (which now can be done with the spin operator alone), then a space translation to the new value $\mathbf{x}'$, then a boost to the $\mathbf{v}$ frame, and finally a time translation forward to t'. This is given by the set:

$$\begin{aligned} G &= (1,0,0,t')(1,\mathbf{v},0,0)(1,0,\mathbf{x}',0)(R,0,0,0)(1,0,-\mathbf{x},0)(1,0,0,-t)\,, \\ &= (1,0,0,t')(1,\mathbf{v},0,0)(1,0,\mathbf{x}',0)(R,0,0,0)(1,0,-\mathbf{x},-t)\,, \\ &= (1,0,0,t')(1,\mathbf{v},0,0)(1,0,\mathbf{x}',0)(R,0,-R(\mathbf{x}),-t)\,, \\ &= (1,0,0,t')(1,\mathbf{v},0,0)(R,0,\mathbf{x}'-R(\mathbf{x}),-t)\,, \\ &= (1,0,0,t')(R,\mathbf{v},\mathbf{x}'-R(\mathbf{x})-\mathbf{v}t,-t)\,, \\ &= (R,\mathbf{v},\mathbf{a},\tau)\,, \end{aligned} \tag{9.173}$$

as required. The combined unitary transformation for the full Galilean group is then given by:

$$U_H(t')\,U_{\mathbf{X}}(\mathbf{v})U_{\mathbf{P}}(\mathbf{x}')\,U_{\mathbf{J}}(R)\,U_{\mathbf{P}}(-\mathbf{x})\,U_H(-t) = e^{ig(\mathbf{x},t)/\hbar}\,U(G)\,. \tag{9.174}$$

The only contribution to the phase comes from between step four and step five in the above. Using Bargmann's theorem, we find:

$$g(\mathbf{x},t) = \frac{1}{2}\,M\,\mathbf{v}\cdot(\mathbf{x}'-R(\mathbf{x})) = \frac{1}{2}\,Mv^2\,t + \frac{1}{2}\,M\,\mathbf{v}\cdot\mathbf{a}\,. \tag{9.175}$$

So

$$\begin{aligned} U(G)\,|\,\mathbf{x},t;sm\,\rangle &= e^{-ig(\mathbf{x},t)/\hbar}\,U_H(t')\,U_{\mathbf{X}}(\mathbf{v})U_{\mathbf{P}}(\mathbf{x}')\,U_{\mathbf{J}}(R)\,U_{\mathbf{P}}(-\mathbf{x})\,U_H(-t)\,|\,\mathbf{x},t;sm\,\rangle \\ &= e^{-ig(\mathbf{x},t)/\hbar}\,U_H(t')\,U_{\mathbf{X}}(\mathbf{v})U_{\mathbf{P}}(\mathbf{x}')\,U_{\mathbf{J}}(R)\,U_{\mathbf{P}}(-\mathbf{x})\,|\,\mathbf{x},0;sm\,\rangle \\ &= e^{-ig(\mathbf{x},t)/\hbar}\,U_H(t')\,U_{\mathbf{X}}(\mathbf{v})U_{\mathbf{P}}(\mathbf{x}')\,U_{\mathbf{J}}(R)\,|\,\mathbf{0},0;sm\,\rangle \\ &= e^{-ig(\mathbf{x},t)/\hbar}\,U_H(t')\,U_{\mathbf{X}}(\mathbf{v})U_{\mathbf{P}}(\mathbf{x}')\,U_{\mathbf{S}}(R)\,|\,\mathbf{0},0;sm\,\rangle \\ &= e^{-ig(\mathbf{x},t)/\hbar}\,U_H(t')\,U_{\mathbf{X}}(\mathbf{v})U_{\mathbf{P}}(\mathbf{x}')\sum_{m'}|\,\mathbf{0},0;sm'\,\rangle\,D^{(s)}_{m',m}(R) \\ &= e^{-ig(\mathbf{x},t)/\hbar}\,U_H(t')\,U_{\mathbf{X}}(\mathbf{v})\sum_{m'}|\,\mathbf{x}',0;sm'\,\rangle\,D^{(s)}_{m',m}(R) \\ &= e^{-ig(\mathbf{x},t)/\hbar}\,U_H(t')\sum_{m'}e^{iM\mathbf{v}\cdot\mathbf{x}'}|\,\mathbf{x}',0;sm'\,\rangle\,D^{(s)}_{m',m}(R) \\ &= e^{if(\mathbf{x},t)/\hbar}\sum_{m'}|\,\mathbf{x}',t';sm'\,\rangle\,D^{(s)}_{m',m}(R) \end{aligned} \tag{9.176}$$

Where we have defined the phase factor $\phi(G)$ by:

$$\begin{aligned} f(\mathbf{x},t) &= M\mathbf{v}\cdot\mathbf{x}' - \chi(G) = M\mathbf{v}\cdot\mathbf{x}' - \frac{1}{2}\,Mv^2\,t - \frac{1}{2}\,M\,\mathbf{v}\cdot\mathbf{a} \\ &= M\mathbf{v}\cdot R(\mathbf{x}) + \frac{1}{2}\,Mv^2\,t + \frac{1}{2}\,M\,\mathbf{v}\cdot\mathbf{a}\,. \end{aligned} \tag{9.177}$$

Inverting Eq. (9.176), we find:

$$U^{\dagger}(G)\,|\,\mathbf{x}',t';sm'\,\rangle = e^{-if(\mathbf{x},t)/\hbar}\sum_{m}|\,\mathbf{x},t;sm\,\rangle\,D^{(s)\,*}_{m',m}(R)\,. \tag{9.178}$$

So that:

$$\psi'_{sm'}(\mathbf{x}',t') = e^{if(\mathbf{x},t)/\hbar}\sum_{m}D^{(s)}_{m',m}(R)\,\psi_{sm}(\mathbf{x},t)\,. \tag{9.179}$$

where $\psi_{sm}(\mathbf{x},t) = \langle\,\mathbf{x},t;sm\,|\,\psi\,\rangle$, and we have put: $|\,\psi'\,\rangle = U(G)\,|\,\psi\,\rangle$. It is important to note here that the phase factor $f(\mathbf{x},t)$ depends on $\mathbf{x}$ and t, as well as the parameters of the Galilean transformation.

Exercise 10. Find the general Galilean transformation of momentum eigenvectors: $|\,\mathbf{p},sm\,\rangle$. Show that the transformed functions $\tilde{\psi}_{sm}(\mathbf{p})$ give the same result as as the Fourier transform of Eq. (9.179).

9.7 Improper transformations

In this section we follow Weinberg[**?**, p. 77]. We first extend the kinds of Galilean transformations we consider to include parity, time reversal, and charge conjugation. The full Galilean transformations are now described by:

$$\mathbf{x}' = rR(\mathbf{x}) + \mathbf{v}t + \mathbf{a}\,,\qquad t' = \kappa t + \tau\,. \tag{9.180}$$

Here $r = \det[\,R\,]$ and κ can have values of ± 1. We still require that lengths are preserved so that R is still orthogonal, and that the rate of passage of time does not dilate or shrink, only the direction of time can be reversed. So the full group, including improper transformations, is now represented by the twelve parameters:

$$G = (\,R,\mathbf{v},\mathbf{a},\tau,r,\kappa\,)\,. \tag{9.181}$$

The full group properties are now stated in the next theorem.

Theorem 24. *The composition rule for the full Galilean group is given by:*

$$\begin{aligned} G'' = G'G &= (\,R',\mathbf{v}',\mathbf{a}',\tau',r',\kappa'\,)\,(\,R,\mathbf{v},\mathbf{a},\tau,r,\kappa\,) \\ &= (\,R'R,\kappa\mathbf{v}' + r'R'(\mathbf{v}),\mathbf{a}' + \mathbf{v}'\tau + r'R(\mathbf{a}),\kappa\tau' + \tau,r'r,\kappa'\kappa,\,) \end{aligned} \tag{9.182}$$

Proof. The proof follows directly from the complete transformation equations (9.180) and left as an exercise. □

9.7.1 Parity

In this section we consider parity transformations (space reversals) of the coordinate system. This is represented by the group elements:

$$G_P = (1,0,0,0,-1,+1)\,. \tag{9.183}$$

We note that $G_P^{-1} = G_P$. So using the rules given in Theorem 24, we find for the combined transformation:

$$\begin{aligned} G' = G_P^{-1}\,G\,G_P &= (1,0,0,0,-1,+1)\,(\,R,\mathbf{v},\mathbf{a},\tau,r,\kappa\,)\,(1,0,0,0,-1,+1) \\ &= (\,R,-\mathbf{v},-\mathbf{a},\tau,r,\kappa\,)\,. \end{aligned} \tag{9.184}$$

The phase factors are zero in this case. So we have:

$$\mathcal{P}^{-1}\,U(G)\,\mathcal{P} = U(G_P^{-1}\,G\,G_P) = U(G')\,. \tag{9.185}$$

Now if we take $r = 1$ and $\kappa = 1$, both G and G' are proper. This means that we can take $G = 1 + \Delta G$, where $\Delta G = (\,\Delta\omega, \Delta\mathbf{v}, \Delta\mathbf{a}, \Delta\tau, 1, 1\,)$. Then $G' = 1 + \Delta G'$, where $\Delta G' = (\,\Delta\omega, -\Delta\mathbf{v}, -\Delta\mathbf{a}, \Delta\tau, 1, 1\,)$. So then $U(1 + \Delta G)$ can be represented by:[5]

$$U(1 + \Delta G) = 1 + \frac{i}{\hbar}\Big\{\Delta\theta\,\hat{\mathbf{n}}\cdot\mathbf{J} + \Delta\mathbf{v}\cdot\mathbf{K} - \Delta\mathbf{a}\cdot\mathbf{P} + \Delta\tau\,H\Big\} + \cdots\,. \tag{9.186}$$

Using this in Eq. (9.185), we find:

$$\begin{aligned}
\mathcal{P}^{-1}\,\mathbf{J}\,\mathcal{P} &= \mathbf{J}\,,\\
\mathcal{P}^{-1}\,\mathbf{K}\,\mathcal{P} &= -\mathbf{K}\,,\\
\mathcal{P}^{-1}\,\mathbf{P}\,\mathcal{P} &= -\mathbf{P}\,,\\
\mathcal{P}^{-1}\,H\,\mathcal{P} &= H\,.
\end{aligned} \tag{9.187}$$

We note that $\mathcal{P}$ is linear and unitary, with eigenvalues of unit magnitude. We also have: $\mathcal{P}^{-1} = \mathcal{P}^\dagger = \mathcal{P}$. We assume that the Casimir invariants M and W remain unchanged by a parity transformation.

Exercise 11. Show that under parity,

$$\mathcal{P}^{-1}\,\mathbf{X}(t)\,\mathcal{P} = -\mathbf{X}(t)\,, \tag{9.188}$$

where $\mathbf{X}(t) = \mathbf{X} + \mathbf{V}\,t$, where $\mathbf{X} = \mathbf{K}/M$ and $\mathbf{V} = \mathbf{P}/M$.

We discuss the action of parity on eigenvectors of angular momentum in Section 21.1.4.

9.7.2 Time reversal

Time reversal is represented by the group elements:

$$G_T = (1, 0, 0, 0, +1, -1)\,, \tag{9.189}$$

with $G_T^{-1} = G_T$. So again using the rules given in Theorem 24, we find for the combined transformation:

$$\begin{aligned}
G' = G_T^{-1}\,G\,G_T &= (1, 0, 0, 0, +1, -1)\,(\,R, \mathbf{v}, \mathbf{a}, \tau, r, \kappa\,)\,(1, 0, 0, 0, +1, -1)\\
&= (\,R, -\mathbf{v}, \mathbf{a}, -\tau, r, \kappa\,)\,.
\end{aligned} \tag{9.190}$$

So we have:

$$\mathcal{T}^{-1}\,U(G)\,\mathcal{T} = U(G_T^{-1}\,G\,G_T) = U(G')\,. \tag{9.191}$$

Again, we take $r = +1$ and $\kappa = +1$, so that $G = 1 + \Delta G$ and $G' = 1 + \Delta G'$, where

$$\begin{aligned}
\Delta G &= \big(\,\Delta\omega, \Delta\mathbf{v}, \Delta\mathbf{a}, \Delta\tau, 1, 1\,\big)\,,\\
\Delta G' &= \big(\,\Delta\omega, -\Delta\mathbf{v}, \Delta\mathbf{a}, -\Delta\tau, 1, 1\,\big)\,,
\end{aligned} \tag{9.192}$$

Both of these transformations are proper. So we can take $U(G)$ and $G(G')$ to be represented by the infinitesimal form of Eq. (9.186). Since we will require $\mathcal{T}$ to be anti-linear and anti-unitary, $\mathcal{T}^{-1} i\,\mathcal{T} = -i$, and, using (9.191), we find:

$$\begin{aligned}
\mathcal{T}^{-1}\,\mathbf{J}\,\mathcal{T} &= -\mathbf{J}\,,\\
\mathcal{T}^{-1}\,\mathbf{K}\,\mathcal{T} &= \mathbf{K}\,,\\
\mathcal{T}^{-1}\,\mathbf{P}\,\mathcal{T} &= -\mathbf{P}\,,\\
\mathcal{T}^{-1}\,H\,\mathcal{T} &= H\,.
\end{aligned} \tag{9.193}$$

We also assume that M and W are *unchanged* by a time-reversal transformation. The eigenvalues of $\mathcal{T}$ are also of unit magnitude. We also have: $\mathcal{T}^{-1} = \mathcal{T}^\dagger = \mathcal{T}$. We discuss time reversal of angular momentum eigenvectors in Section 21.1.4.

[5] We do not use the extended group in this discussion.

Exercise 12. Show that under time reversal,

$$\mathcal{T}^{-1}\,\mathbf{X}(t)\,\mathcal{T} = \mathbf{X}(-t)\,, \tag{9.194}$$

where $\mathbf{X}(t) = \mathbf{X} + \mathbf{V}\,t$, where $\mathbf{X} = \mathbf{K}/M$ and $\mathbf{V} = \mathbf{P}/M$.

For combined parity and time-reversal transformations, we find:

$$\begin{aligned}(\mathcal{PT})^{-1}\,\mathbf{J}\,(\mathcal{PT}) &= -\mathbf{J}\,,\\ (\mathcal{PT})^{-1}\,\mathbf{K}\,(\mathcal{PT}) &= -\mathbf{K}\,,\\ (\mathcal{PT})^{-1}\,\mathbf{P}\,(\mathcal{PT}) &= \mathbf{P}\,,\\ (\mathcal{PT})^{-1}\,H\,(\mathcal{PT}) &= H\,.\end{aligned} \tag{9.195}$$

9.7.3 Charge conjugation

The charge conjugation operator $\mathcal{C}$ changes particles into antiparticles. This is not a space-time symmetry, but one that reverses the sign of the mass and spin. That is, we assume that:

$$\mathcal{C}^{-1}\,M\,\mathcal{C} = -M\,, \qquad \mathcal{C}^{-1}\,\mathbf{S}\,\mathcal{C} = -\mathbf{S}\,. \tag{9.196}$$

In addition, we take $\mathcal{C}$ to be linear and unitary, and:

$$\begin{aligned}\mathcal{C}^{-1}\,\mathbf{J}\,\mathcal{C} &= -\mathbf{J}\,,\\ \mathcal{C}^{-1}\,\mathbf{K}\,\mathcal{C} &= -\mathbf{K}\,,\\ \mathcal{C}^{-1}\,\mathbf{P}\,\mathcal{C} &= \mathbf{P}\,,\\ \mathcal{C}^{-1}\,H\,\mathcal{C} &= H\,.\end{aligned} \tag{9.197}$$

The eigenvalues of $\mathcal{C}$ are again of unit magnitude. If we define $\mathbf{X} = \mathbf{K}/M$, and $\mathbf{V} = \mathbf{P}/M$, then this means that

$$\begin{aligned}\mathcal{C}^{-1}\,\mathbf{X}\,\mathcal{C} &= \mathbf{X}\,,\\ \mathcal{C}^{-1}\,\mathbf{V}\,\mathcal{C} &= -\mathbf{V}\,,\end{aligned} \tag{9.198}$$

So we have the following theorem:

Theorem 25 ($\mathcal{PTC}$)**.** *From Eqs. (9.195) and (9.197), the combined* ($\mathcal{PTC}$) *operation when acting on the generators of the Galilean transformation, leaves the generators unchanged:*

$$\begin{aligned}(\mathcal{PTC})^{-1}\,\mathbf{J}\,(\mathcal{PTC}) &= \mathbf{J}\,,\\ (\mathcal{PTC})^{-1}\,\mathbf{K}\,(\mathcal{PTC}) &= \mathbf{K}\,,\\ (\mathcal{PTC})^{-1}\,\mathbf{P}\,(\mathcal{PTC}) &= \mathbf{P}\,,\\ (\mathcal{PTC})^{-1}\,H\,(\mathcal{PTC}) &= H\,.\end{aligned} \tag{9.199}$$

That is, the generators are invariant under ($\mathcal{PTC}$).

Exercise 13. Show that under charge conjugation,

$$\mathcal{C}^{-1}\,\mathbf{X}(t)\,\mathcal{C} = \mathbf{X}(-t)\,, \tag{9.200}$$

where $\mathbf{X}(t) = \mathbf{X} + \mathbf{V}\,t$, with $\mathbf{X} = \mathbf{K}/M$ and $\mathbf{V} = \mathbf{P}/M$. So when acting on the equation of motion of $\mathbf{X}(t)$, charge conjugation has the same effect as time reversal. We can interpret this as meaning that in non-relativistic physics, we can think of an antiparticle as a negative mass particle moving backwards in time.

Let us be precise. If $|\,\psi\,\rangle$ represents a single particle state, then $|\,\psi_c\,\rangle = \mathcal{C}\,|\,\psi\,\rangle$ is the charge conjugate state. Ignoring spin for the moment, if $|\,m_0, w_0, E_0; \mathbf{x}, t\,\rangle$ are eigenstates of $\mathbf{X}(t)$ and M with positive eigenvalues $m = m_0 > 0$, $w = w_0 > 0$ and $E = E_0 > 0$, then

$$\mathcal{C}\,|\,m_0, w_0, E_0; \mathbf{x}, t\,\rangle = |\,-m_0, -w_0, -E_0; \mathbf{x}, t\,\rangle\,, \tag{9.201}$$

is an eigenvector $\mathbf{X}(t)$, M, W, and H with negative eigenvalues $m = -m_0 < 0$, $w = -m_0 < 0$, and $E = -E_0 < 0$. So the charge conjugate wave function with m, w, and E all positive:

$$\begin{aligned}\psi_c(m_0, w_0, E_0; \mathbf{x}, t) = \langle\, m_0, w_0, E_0; \mathbf{x}, t\,|\,\psi_c\,\rangle &= \langle\, m_0, w_0, E_0; \mathbf{x}, t\,|\,\mathcal{C}\,|\,\psi\,\rangle \\ &= \langle\, -m_0, -w_0, -E_0; \mathbf{x}, t\,|\,\psi\,\rangle = \psi(-m_0, -w_0, -E_0; \mathbf{x}, t)\,,\end{aligned} \tag{9.202}$$

is the same as the wave function with m_0, w_0, and E_0 negative. We will study single particle wave functions in the next chapter. Charge conjugate symmetry says that, in priciple, we cannot tell the difference between a world consisting of particles or a world consisting of antiparticles.

9.8 Scale and conformal transformations

Scale transformations are changes in the measures of length and time. An interesting question is if there are ways to determine a length or time scale in absolute terms, or are these just arbitrary measures. If there are no physical systems that can set these scales, we say that the fundamental forces in Nature must be scale invariant. Conformal invariance is a combined space-time expansion of the measures of length and time, and generalizes scale changes. We discuss these additional space-time symmetries in the next two sections.

9.8.1 Scale transformations

Scale transformations are of the form:

$$x_i' = \alpha\, x_i\,, \qquad t' = \beta\, t\,. \tag{9.203}$$

We require, in particular, that if $\psi(\mathbf{x}, t)$ satisfies Schrödinger's equation with $w = 0$ for a spinless free particle in Σ, then $\psi'(\mathbf{x}', t')$ satisfies Schrödinger's equation in Σ'. Probability must remain the same, so we require that

$$|\,\psi'(\mathbf{x}', t')\,|^2\,\mathrm{d}^3x' = |\,\psi(\mathbf{x}, t)\,|^2\,\mathrm{d}^3x\,. \tag{9.204}$$

With this observation, it is easy to prove the following theorem.

Theorem 26. *Under scale transformations $\mathbf{x}' = \alpha\mathbf{x}$ and $t' = \beta t$, spinless scalar solutions of Schrödinger's equation transform according to:*

$$\psi'(\mathbf{x}', t') = \alpha^{-3/2}\, e^{ig(\mathbf{x},t)/\hbar}\,\psi(\mathbf{x}, t)\,. \tag{9.205}$$

with $\beta = \alpha^2$ and $g(\mathbf{x}, t) = C$, a constant phase.

Exercise 14. Prove Theorem 26.

We put $\alpha = e^s$ and then $\beta = e^{2s}$, so that infinitesimal scale transformations become:

$$\Delta\mathbf{x} = \Delta s\,\mathbf{x}\,, \qquad \Delta t = 2\,\Delta s\, t\,. \tag{9.206}$$

We now follow our work in example 27 to find a differential representation of the scale generator D. Using Eq. (9.205), infinitesimal scale changes of scalar functions are given by:

$$\begin{aligned}\psi'(\mathbf{x}', t') &= e^{-3\Delta s/2}\,\psi(\mathbf{x}' - \Delta\mathbf{x}, t' - \Delta t) \\ &= \left\{\,1 - 3\Delta s/2 + \cdots\,\right\}\left\{\,1 - \Delta s\,\mathbf{x}\cdot\boldsymbol{\nabla} - \Delta s\, 2\, t\,\partial_t + \cdots\,\right\}\psi(\mathbf{x}', t') \\ &= \left\{\,1 - \Delta s\left\{\,3/2 + \mathbf{x}\cdot\boldsymbol{\nabla} + 2\,t\,\partial_t\,\right\} + \cdots\,\right\}\psi(\mathbf{x}', t')\end{aligned} \tag{9.207}$$

The **dilation** generator D is defined by:

$$\Delta\psi(\mathbf{x},t) = \psi'(\mathbf{x},t) - \psi(\mathbf{x},t) = -i\,\Delta s\,D\,\psi(\mathbf{x},t)\,, \tag{9.208}$$

from which we find:

$$D = -\frac{3}{2}\,i + \frac{1}{i}\,\mathbf{x}\cdot\boldsymbol{\nabla} - 2i\,t\,\partial_t = -\frac{3}{2}\,i + \mathbf{x}\cdot\mathbf{P} - 2\,t\,H\,. \tag{9.209}$$

We can drop the factor of $-3i/2$ since this produces only a constant phase. Using the differential representations in Eqs. (9.13), we find the commutation relations for D:

$$[\,D,P_i\,] = iP_i\,, \qquad [\,D,H\,] = 2i\,H\,, \qquad [\,D,K_i\,] = -iK_i\,, \tag{9.210}$$

and commutes with J_i. D also commutes with M, but we note that the first Casimir operator $W = H = P^2/2M$ does not commute with D. In fact, we find:

$$[\,D,W\,] = 2i\,W\,. \tag{9.211}$$

So the internal energy W breaks scale symmetry.

9.8.2 Conformal transformations

Conformal transformations are of the form:

$$x_i' = \frac{x_i}{1-ct}\,, \qquad t' = \frac{t}{1-ct}\,, \tag{9.212}$$

where c has units of reciprocal time (not velocity!) and can be positive or negative. Note that $1/t' = 1/t - c$.

For a scalar spin zero free particle satisfying Schrödinger's equation, probability is again conserved according to (9.204), and we find the following result for conformal transformations:

Theorem 27. *Under scale transformations $\mathbf{x}' = \alpha\mathbf{x}$ and $t' = \beta t$, spinless scalar solutions of Schrödinger's equation transform according to:*

$$\psi'(\mathbf{x}',t') = (1-ct)^{3/2}\,e^{ig(\mathbf{x},t)/\hbar}\,\psi(\mathbf{x},t)\,. \tag{9.213}$$

where

$$g(\mathbf{x},t) = \frac{1}{2}\,\frac{mc\,x^2}{1-ct}\,. \tag{9.214}$$

Exercise 15. Prove Theorem 27. For this, it is useful to note that:

$$\boldsymbol{\nabla}' = (1-ct)\,\boldsymbol{\nabla}\,, \qquad \partial_t' = (1-ct)^2\,\partial_t - c(1-ct)\,\mathbf{x}\cdot\boldsymbol{\nabla}\,. \tag{9.215}$$

and that:

$$\frac{\hbar}{i}\boldsymbol{\nabla}\left[\,e^{ig(\mathbf{x},t)/\hbar}\,\psi(\mathbf{x},t)\right] = e^{ig(\mathbf{x},t)/\hbar}\left[\,\frac{\hbar}{i}\boldsymbol{\nabla} + (\boldsymbol{\nabla}g(\mathbf{x},t))\,\right]\psi(\mathbf{x},t)\,. \tag{9.216}$$

Infinitesimal conformal transformations are given by:

$$\Delta\mathbf{x} = \Delta c\,t\,\mathbf{x}\,, \qquad \Delta t = \Delta c\,t^2\,. \tag{9.217}$$

So from Eq. (9.213), infinitesimal conformal transformations of scalar functions are given by:

$$\psi'(\mathbf{x}',t') = (1-t\,\Delta c)^{3/2}\,e^{i\Delta g(\mathbf{x}',t')/\hbar}\,\psi(\mathbf{x}'-\Delta\mathbf{x},t'-\Delta t)\,, \tag{9.218}$$

where

$$\Delta g(\mathbf{x}',t') = \frac{1}{2}\,mx^2\,\Delta c\,. \tag{9.219}$$

So

$$\begin{aligned}\psi'(\mathbf{x}',t') &= \Big\{1-\frac{3}{2}\,t\,\Delta c+\cdots\Big\}\Big\{1+\frac{\hbar}{2i}\,mx^2\,\Delta c+\cdots\Big\}\\ &\quad\times\Big\{1-\Delta c\,t\mathbf{x}\cdot\boldsymbol{\nabla}-\Delta c\,t^2\,\partial_t+\cdots\Big\}\psi(\mathbf{x}',t')\\ &=\Big\{1+\Delta c\Big\{-\frac{3}{2}\,t+\frac{\hbar}{2i}\,mx^2-t\mathbf{x}\cdot\boldsymbol{\nabla}-t^2\,\partial_t\Big\}+\cdots\Big\}\psi(\mathbf{x}',t')\,,\end{aligned}\tag{9.220}$$

The **conformal** generator C is defined by:

$$\Delta\psi(\mathbf{x},t)=\psi'(\mathbf{x},t)-\psi(\mathbf{x},t)=i\,\Delta c\,C\,\psi(\mathbf{x},t)\,,\tag{9.221}$$

from which we find:

$$\begin{aligned}C&=\frac{3i}{2}\,t-\frac{\hbar}{2}\,mx^2+\frac{t}{i}\,\mathbf{x}\cdot\boldsymbol{\nabla}-i\,t^2\,\partial_t\\ &=\frac{3i}{2}\,t-\frac{\hbar}{2}\,mx^2+t\,\mathbf{x}\cdot\mathbf{P}-t^2\,H\\ &=\frac{3i}{2}\,t-\frac{\hbar}{2}\,mx^2+t\,D+t^2\,H\,.\end{aligned}\tag{9.222}$$

We find the following commutation relations for C:

$$[\,C,H\,]=-i\,D\,,\qquad[\,C,D\,]=-2i\,C\,,\tag{9.223}$$

and commutes with all other operators. Note that scale and conformal transformations do *not* commute. So if we put:

$$G_1=\frac{1}{2}\,(H+C)\,,\qquad G_2=\frac{1}{2}\,(H-C)\,,\qquad G_3=\frac{1}{2}\,D\,,\tag{9.224}$$

we find that $\mathbf{G}$ satisfies a $O(2,1)$ algebra:

$$[\,G_1,G_2\,]=-i\,G_3\,,\qquad[\,G_1,G_3\,]=i\,G_2\,,\qquad[\,G_2,G_3\,]=i\,G_1\,.\tag{9.225}$$

Since $[\,G_i,J_j\,]=0$, the group structure of the extended group has $O(3)\times O(2,1)$ symmetry.

9.9 The Schrödinger group

The extension of the Galilean group to include scale and conformal transformations is called the **Schrödinger** or non-relativistic conformal group, which we write as $\mathcal{S}$. We consider combined scale and conformal transformations of the following form:

$$\mathbf{x}'=\frac{R(\mathbf{x})+\mathbf{v}t+\mathbf{a}}{\gamma t+\delta}\,,\qquad t'=\frac{\alpha t+\beta}{\gamma t+\delta}\,,\qquad\alpha\delta-\beta\gamma=1\,.\tag{9.226}$$

Here α, β, γ, and δ are real parameters, only three of which are independent. This transformation contains both scale and conformal transformations as special interrelated cases. The group elements now consist of twelve independent parameters, but it is useful to write them in terms of thirteen parameters with one constraint: $S=(R,\mathbf{v},\mathbf{a},\alpha,\beta,\gamma,\delta)$. The extended transformation *is* a group. The group multiplication properties are contained in the next theorem:

Theorem 28. *The multiplication law for the Schrödinger group is given by:*

$$\begin{aligned}S''=S'S&=(R',\mathbf{v}',\mathbf{a}',\alpha',\beta',\gamma',\delta')\,(R,\mathbf{v},\mathbf{a},\alpha,\beta,\gamma,\delta)\\ &=(\,R'R,R'(\mathbf{v})+\alpha\mathbf{v}'+\gamma\mathbf{a}',R'(\mathbf{a})+\beta\mathbf{v}'+\delta\mathbf{a}',\\ &\qquad\alpha'\alpha+\beta'\gamma,\alpha'\beta+\beta'\delta,\gamma'\alpha+\delta'\gamma,\gamma'\beta+\delta'\delta\,)\,.\end{aligned}\tag{9.227}$$

A faithful five-dimensional matrix representation is given by:

$$\underline{S} = \begin{pmatrix} R & \mathbf{v} & \mathbf{a} \\ 0 & \alpha & \beta \\ 0 & \gamma & \delta \end{pmatrix} , \qquad \underline{S}'' = \underline{S}'\underline{S} , \tag{9.228}$$

which preserves the determinant relation: $\det[\, S\,] = \alpha\delta - \beta\gamma = 1$. The unit element is $1 = (1, 0, 0, 1, 0, 0, 1)$ and the inverse element is:

$$S^{-1} = (R^{-1}, -\delta R^{-1}(\mathbf{v}) + \gamma R^{-1}(\mathbf{a}), -\alpha R^{-1}(\mathbf{a}) + \beta R^{-1}(\mathbf{v}), \delta, -\beta, -\gamma, \alpha\,) . \tag{9.229}$$

For infinitesimal transformations, it is useful to write:

$$\begin{aligned} \alpha &= 1 + \Delta s + \cdots , \\ \beta &= \Delta\tau + \cdots , \\ \gamma &= -\Delta c + \cdots , \\ \delta &= 1 - \Delta s + \cdots , \end{aligned} \tag{9.230}$$

so that

$$\alpha\delta - \beta\gamma = (\,1 + \Delta s + \cdots\,)\,(\,1 - \Delta s + \cdots\,) - (\,\Delta\tau + \cdots\,)\,(\,-\Delta c + \cdots\,) = 1 + O(\Delta^2) , \tag{9.231}$$

as required. $\Delta\tau$, Δs, and Δc are now independent variations. So the unitary transformation transformation for infinitesimal transformations is now written as:

$$U(1 + \Delta S) = 1 + \frac{i}{\hbar}\left\{\Delta\theta\,\hat{\mathbf{n}}\cdot\mathbf{J} + \Delta\mathbf{v}\cdot\mathbf{K} - \Delta\mathbf{a}\cdot\mathbf{P} + \Delta\tau\,H + \Delta s\,D - \Delta c\,C\right\} + \cdots , \tag{9.232}$$

in terms of the twelve generators $\mathbf{J}$, $\mathbf{K}$, $\mathbf{P}$, H, D, and C.

References

[1] E. P. Wigner, *Gruppentheorie und ihre Anwendung auf dei Quantenmechanic der Atomspektren* (Braunschweig, Berlin, 1931). English translation: Academic Press, Inc, New York, 1959.

[2] V. Bargmann, "On unitary ray representations of continuous groups," Ann. Math. **59**, 1 (1954).

[3] J.-M. Levy-Leblond, "Galilei group and nonrelativistic quantum mechanics," J. Math. Phys. **4**, 776 (1963).

[4] J.-M. Levy-Leblond, "Galilean quantum field theories and a ghostless Lee model," Commun. Math. Phys. **4**, 157 (1967).

[5] J.-M. Levy-Leblond, "Nonrelativistic particles and wave equations," Commun. Math. Phys. **6**, 286 (1967).

[6] J.-M. Levy-Leblond, "Galilei group and galilean invariance," in E. M. Loebl (editor), "Group theory and its applications," volume II, pages 222–296 (Academic Press, New York, NY, 1971).

Chapter 10

Wave equations

In this chapter, we discuss wave equations for single free particles. We first discuss wave equations for a single free particle of mass $M \neq 0$ and for a fixed value of $w = 0$. We find wave equations for scalar ($s = 0$), spinor ($s = 1/2$), and vector ($s = 1$) particles.

10.1 Scalars

For scalar particles with $s = 0$, let us define time dependent wave single particle $(+)$ and antiparticle $(-)$ wave funtions for $m = \pm m_0$ and $w = \pm w_0$ by:

$$\psi^{(\pm)}(\mathbf{x}, t) = \langle\, \pm m_0, \pm w_0; \mathbf{x}, t \,|\, \psi \,\rangle \,, \tag{10.1}$$

where $m_0 > 0$. From the first Casimir invariant, Eq. (9.80), where

$$H = W + \frac{P^2}{2M} \,, \tag{10.2}$$

and from the time-displacement operator Eq. (9.129), and the coordinate representation of the momentum operator, Eq. (9.146), we find Schrödinger's wave equation for a spinless particle:

$$i\hbar \frac{\partial}{\partial t} \psi^{(\pm)}(\mathbf{x}, t) = \left\{ \mp \frac{\hbar^2}{2m_0} \nabla^2 \pm w_0 \right\} \psi^{(\pm)}(\mathbf{x}, t) \,. \tag{10.3}$$

This equation obeys a probability conservation equation, given by:

$$\frac{\partial \rho^{(\pm)}(\mathbf{x}, t)}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{j}^{(\pm)}(\mathbf{x}, t) = 0 \,, \tag{10.4}$$

where

$$\begin{aligned} \rho^{(\pm)}(\mathbf{x}, t) &= |\psi^{(\pm)}(\mathbf{x}, t)|^2 \,, \\ \mathbf{j}^{(\pm)}(\mathbf{x}, t) &= \pm \frac{\hbar}{2m_0\, i} \left[\, \psi^{(\pm)\,*}(\mathbf{x}, t) \left(\boldsymbol{\nabla}\, \psi^{(\pm)}(\mathbf{x}, t) \right) - \left(\boldsymbol{\nabla}\, \psi^{(\pm)\,*}(\mathbf{x}, t) \right) \psi^{(\pm)}(\mathbf{x}, t) \,\right] . \end{aligned} \tag{10.5}$$

We interpret $|\psi^{(\pm)}(\mathbf{x}, t)|^2$ as the probability of finding the particle at point $\mathbf{x}$ at time t.

Now the particle and antiparticle solutions are related by:

$$\psi^{(\pm)}(\mathbf{x}, t) = \mathcal{K}[\, \psi^{(\mp)}(\mathbf{x}, t) \,] = \psi^{(\mp)\,*}(\mathbf{x}, t) \,, \tag{10.6}$$

where $\mathcal{K}$ is a complex conjugation operator.

Exercise 16. Show that $\mathcal{K}$ is an anti-linear anti-unitary operator with eigenvalues of unit magnitude.

General solutions for particles and antiparticles of (10.3) can be given as Fourier transforms:

$$\begin{aligned}\psi^{(+)}(\mathbf{x},t) &= \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, a^{(+)}_{\mathbf{k}}\, e^{+i(\mathbf{k}\cdot\mathbf{x}-E_kt)}\,,\\ \psi^{(-)}(\mathbf{x},t) &= \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, a^{(-)}_{\mathbf{k}}\, e^{+i(\mathbf{k}\cdot\mathbf{x}+E_kt)} = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, a^{(-)}_{-\mathbf{k}}\, e^{-i(\mathbf{k}\cdot\mathbf{x}-E_kt)}\,,\end{aligned} \tag{10.7}$$

where $E_k = \hbar k^2/(2m_0) + w_0$ in all integrals. In the integral in the last line, we have put $\mathbf{k} \to -\mathbf{k}$. We first note that $\psi^{(-)}(-m_0,-w_0;\mathbf{x},t) = \psi^{(+)}(+m_0,+w_0;\mathbf{x},t)$, as required.

Under a Galilean transformation of space-time, a scalar wave function transforms like:

$$\psi^{(\pm)\prime}(\mathbf{x}',t') = e^{\pm if(\mathbf{x},t)/\hbar}\,\psi^{(\pm)}(\mathbf{x},t)\,, \tag{10.8}$$

where

$$f(\mathbf{x},t) = m_0\mathbf{v}\cdot R(\mathbf{x}) + \frac{1}{2}\,m_0v^2\,t + \frac{1}{2}\,m_0\,\mathbf{v}\cdot\mathbf{a}\,. \tag{10.9}$$

We see from this that particle wave functions transform differently than antiparticle wave functions. This difference in transformation properties is called Bargmann's **superselection rule** and means that we cannot add particle wave functions to antiparticle wave functions and maintain Galilean invariance of the result. The best we can do is construct a two-component wave function $\Psi(\mathbf{x},t)$ by the definition:

$$\Psi(\mathbf{x},t) = \begin{pmatrix}\psi^{(+)}(\mathbf{x},t)\\ \psi^{(-)}(\mathbf{x},t)\end{pmatrix}\,, \tag{10.10}$$

which transforms according to:

$$\Psi'(\mathbf{x}',t') = S(\mathbf{x},t)\,\Psi(\mathbf{x},t)\,,\qquad \text{where}\qquad S(\mathbf{x},t) = \begin{pmatrix}e^{+if(\mathbf{x},t)/\hbar} & 0\\ 0 & e^{-if(\mathbf{x},t)/\hbar}\end{pmatrix}\,. \tag{10.11}$$

Exercise 17. Show directly by differentiation that if $\psi^{(+)}(\mathbf{x},t)$ satisfies Schrödinger's equation in frame Σ:

$$i\hbar\,\frac{\partial}{\partial t}\,\psi^{(+)}(\mathbf{x},t) = \left\{-\frac{\hbar^2}{2m_0}\nabla^2 + w_0\right\}\psi^{(+)}(\mathbf{x},t)\,. \tag{10.12}$$

then $\psi^{(+)\prime}(\mathbf{x}',t')$, given by Eq. (10.8), satisfies Schrödinger's equation in frame Σ':

$$i\hbar\,\frac{\partial}{\partial t'}\,\psi^{(+)\prime}(\mathbf{x}',t') = \left\{-\frac{\hbar^2}{2m_0}\nabla'^2 + w_0\right\}\psi^{(+)\prime}(\mathbf{x}',t')\,. \tag{10.13}$$

10.2 Spinors

In this section, we derive wave equations for spin 1/2 particles and antiparticles.

10.2.1 Spinor particles

For spin 1/2 particles, the time dependent wave functions can be written as two-component column matrices (called **spinors**). However, it is useful to introduce *four*-component column spinors $\psi^{(+)}(\mathbf{x},t)$, which we will call **Pauli spinors**,[1] consisting of a pair of two-component spinors $\phi^{(+)}(\mathbf{x},t)$ and $\chi^{(+)}(\mathbf{x},t)$, for reasons

[1] As opposed to Dirac spinors in the relativistic case.

that will be come apparent later. These are defined by the matrices:

$$\psi^{(+)}(\mathbf{x},t)=\begin{pmatrix}\phi^{(+)}(\mathbf{x},t)\\ \chi^{(+)}(\mathbf{x},t)\end{pmatrix}\,,\qquad \begin{aligned}\phi^{(+)}(\mathbf{x},t)&=\begin{pmatrix}\phi^{(+)}_{+1/2}(\mathbf{x},t)\\ \phi^{(+)}_{-1/2}(\mathbf{x},t)\end{pmatrix}\,,\qquad \phi^{(+)}_{sm}(\mathbf{x},t)=\langle\,\mathbf{x},t;s,m\,|\,\phi\,\rangle\,,\\ \chi^{(+)}(\mathbf{x},t)&=\begin{pmatrix}\chi^{(+)}_{+1/2}(\mathbf{x},t)\\ \chi^{(+)}_{-1/2}(\mathbf{x},t)\end{pmatrix}\,,\qquad \chi^{(+)}_{sm}(\mathbf{x},t)=\langle\,\mathbf{x},t;s,m\,|\,\chi\,\rangle\,,\end{aligned}$$

with $s=1/2$ and $m=\pm1/2$. The $(+)$ sign indicates that $m=m_0>0$. The wave equation can be written as a *second* order differential equation, independent of spin, exactly as Eq. (10.3). However, we shall see that there is some advantage of writing this equation as two coupled first order differential equations. We start by writing a 4×4 matrix equation:

$$\begin{pmatrix}i\hbar\,\partial/\partial t & -\hbar\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}/i\\ -\hbar\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}/i & 2m_0\end{pmatrix}\begin{pmatrix}\phi^{(+)}(\mathbf{x},t)\\ \chi^{(+)}(\mathbf{x},t)\end{pmatrix}=0\,,\tag{10.14}$$

which couples the two-component Pauli spinors $\phi^{(+)}(\mathbf{x},t)$ and $\chi^{(+)}(\mathbf{x},t)$. The solution of Eq. (10.14) is simple. It is given by:

$$\chi^{(+)}(\mathbf{x},t)=\frac{1}{2m_0}\,\frac{\hbar}{i}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,\phi^{(+)}(\mathbf{x},t)\,,\qquad i\hbar\,\frac{\partial}{\partial t}\,\phi^{(+)}(\mathbf{x},t)=\frac{\hbar}{i}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,\chi^{(+)}(\mathbf{x},t)\,,\tag{10.15}$$

which leads to the usual *second* order Schrödinger wave equation for the spinor $\phi(\mathbf{x},t)$:

$$-\frac{\hbar^2}{2m_0}\,(\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,)^2\,\phi^{(+)}(\mathbf{x},t)=-\frac{\hbar^2}{2m_0}\,I\,\nabla^2\,\phi^{(+)}(\mathbf{x},t)=i\hbar\,\frac{\partial}{\partial t}\,\phi^{(+)}(\mathbf{x},t)\,.\tag{10.16}$$

where we have used Eq. (15.4) in Appendix **??**. Eq. (10.14) is called the **Pauli** equation, and we will use it to describe spin 1/2 particles.

From the Pauli equation and its adjoint, we find that the probability density obeys a conservation equation given by:

$$\frac{\partial\rho^{(+)}(\mathbf{x},t)}{\partial t}+\boldsymbol{\nabla}\cdot\mathbf{j}^{(+)}(\mathbf{x},t)=0\,,\tag{10.17}$$

where

$$\begin{aligned}\rho^{(+)}(\mathbf{x},t)&=\phi^{(+)\,\dagger}(\mathbf{x},t)\,\phi^{(+)}(\mathbf{x},t)=\psi^{(+)\,\dagger}(\mathbf{x},t)\,P^{(+)}\,\psi^{(+)}(\mathbf{x},t)\,,\\ \mathbf{j}^{(+)}(\mathbf{x},t)&=\phi^{(+)\,\dagger}(\mathbf{x},t)\,\boldsymbol{\sigma}\,\chi^{(+)}(\mathbf{x},t)+\chi^{(+)\,\dagger}(\mathbf{x},t)\,\boldsymbol{\sigma}\,\phi^{(+)}(\mathbf{x},t)\\ &=\frac{\hbar}{2m_0\,i}\,\big[\,\phi^{(+)\,\dagger}(\mathbf{x},t)\,(\,\boldsymbol{\nabla}\,\phi^{(+)}(\mathbf{x},t)\,)-(\,\boldsymbol{\nabla}\,\phi^{(+)\,\dagger}(\mathbf{x},t)\,)\,\phi^{(+)}(\mathbf{x},t)\,\big]+\boldsymbol{\nabla}\times\mathbf{s}^{(+)}(\mathbf{x},t)\,,\end{aligned}\tag{10.18}$$

where

$$P^{(+)}=\begin{pmatrix}1&0\\0&0\end{pmatrix}\,,\qquad \mathbf{s}^{(+)}(\mathbf{x},t)=\frac{\hbar}{2m_0}\,\big[\,\phi^{(+)\,\dagger}(\mathbf{x},t)\,\boldsymbol{\sigma}\,\phi^{(+)}(\mathbf{x},t)\,\big]\,,\tag{10.19}$$

is the spin probability density. We will see later that we can interpret $\boldsymbol{\mu}(\mathbf{x},t)=q\,\mathbf{s}(\mathbf{x},t)$, where q is the electronic charge, as the magnetic moment of the particle.

Exercise 18. Establish Eqs. (10.17) and (10.18) by using the Pauli equation and the algebra of the Pauli matrices given in Appendix **??**.

Definition 24. Now let us introduce some notation. We put $\tilde{E}=i\hbar\,\partial/\partial t$ and $\tilde{\mathbf{p}}=\hbar\,\boldsymbol{\nabla}/i$, and let us define the differential operator $\mathcal{D}^{(+)}(\mathbf{x},t)$ by:

$$\mathcal{D}^{(+)}(\mathbf{x},t)=\begin{pmatrix}\tilde{E} & -\boldsymbol{\sigma}\cdot\tilde{\mathbf{p}}\\ -\boldsymbol{\sigma}\cdot\tilde{\mathbf{p}} & 2m_0\end{pmatrix}\,.\tag{10.20}$$

Then Eq. (10.14) becomes: $\mathcal{D}^{(+)}(\mathbf{x},t)\,\psi^{(+)}(\mathbf{x},t)=0$.

Next we prove the following theorem, which establishes the properties of the Pauli equation and Pauli spinors under general Galilean transformations.

Theorem 29. *We show here that:*

$$[\,\Lambda^{(+)}(R,\mathbf{v})\,]^\dagger\, e^{-if(\mathbf{x},t)/\hbar}\,\mathcal{D}^{(+)}(\mathbf{x}',t')\, e^{if(\mathbf{x},t)/\hbar}\,\Lambda^{(+)}(R,\mathbf{v}) = \mathcal{D}^{(+)}(\mathbf{x},t)\,, \tag{10.21}$$

where $\mathbf{x}' = R(\mathbf{x}) + \mathbf{v}t + \mathbf{a}$ *and* $t' = t + \tau$*, and from Eq.* (10.9)*,*

$$f(\mathbf{x},t) = m_0\mathbf{v}\cdot R(\mathbf{x}) + \tfrac{1}{2}\,m_0 v^2\,t + \tfrac{1}{2}\,m_0\,\mathbf{v}\cdot\mathbf{a}\,. \tag{10.22}$$

The $\Lambda^{(\pm)}(R,\mathbf{v})$ *matrices are defined in Eq.* (9.108)*.*

Proof. From (9.3), we find:

$$\frac{\partial}{\partial x_i} = \frac{\partial x'_j}{\partial x_i}\,\frac{\partial}{\partial x'_j} = R_{ji}\,\frac{\partial}{\partial x'_j}\,, \qquad \text{or} \qquad \boldsymbol{\nabla}' = R(\boldsymbol{\nabla})\,,$$
$$\frac{\partial}{\partial t} = \frac{\partial t'}{\partial t} = \frac{\partial}{\partial t'} + v_i\,\frac{\partial}{\partial x'_j}\,, \qquad \text{or} \qquad \frac{\partial}{\partial t'} = \frac{\partial}{\partial t} - \mathbf{v}\cdot R(\boldsymbol{\nabla})\,.$$

In terms of our notation for the differential operators $\tilde{E}$ and $\tilde{\mathbf{p}}$, we have:

$$\tilde{E}' = \tilde{E} + \mathbf{v}\cdot R(\tilde{\mathbf{p}})\,, \qquad \tilde{\mathbf{p}}' = R(\tilde{\mathbf{p}})\,, \tag{10.23}$$

So we find:

$$e^{-if(\mathbf{x},t)/\hbar}\,\tilde{\mathbf{p}}'\,e^{if(\mathbf{x},t)/\hbar} = e^{-if(\mathbf{x},t)/\hbar}\,R(\tilde{\mathbf{p}})\,e^{if(\mathbf{x},t)/\hbar} = R(\tilde{\mathbf{p}}) + m_0\mathbf{v}\,,$$

and

$$\begin{aligned} e^{-if(\mathbf{x},t)/\hbar}\,\tilde{E}'\,e^{if(\mathbf{x},t)/\hbar} &= e^{-if(\mathbf{x},t)/\hbar}\,(\,\tilde{E} + \mathbf{v}\cdot R(\tilde{\mathbf{p}})\,e^{if(\mathbf{x},t)/\hbar} \\ &= \tilde{E} - \tfrac{1}{2}m_0v^2 + \mathbf{v}\cdot R(\tilde{\mathbf{p}}) + m_0\,v^2 = \tilde{E} + \tfrac{1}{2}m_0v^2 + \mathbf{v}\cdot R(\tilde{\mathbf{p}})\,. \end{aligned}$$

So we find:

$$e^{-if(\mathbf{x},t)/\hbar}\,\mathcal{D}^{(+)}(\mathbf{x}',t')\,e^{if(\mathbf{x},t)/\hbar} = \begin{pmatrix} \tilde{E} + \frac{1}{2}m_0v^2 + \mathbf{v}\cdot R(\tilde{\mathbf{p}})\,, & -\boldsymbol{\sigma}\cdot[R(\tilde{\mathbf{p}}) + m_0\mathbf{v}] \\ -\boldsymbol{\sigma}\cdot[R(\tilde{\mathbf{p}}) + m_0\mathbf{v})]\,, & 2m_0 \end{pmatrix}.$$

From (9.108),

$$\Lambda^{(+)}(R,\mathbf{v}) = V^{(+)}(\mathbf{v})\,U(R)\,,$$

and from Eqs. (9.104) and (9.105), we find:

$$\begin{aligned} &[\,\Lambda^{(+)}(R,\mathbf{v})\,]^\dagger\, e^{-if(\mathbf{x},t)/\hbar}\,\mathcal{D}^{(+)}(\mathbf{x}',t')\,e^{if(\mathbf{x},t)/\hbar}\,\Lambda^{(+)}(R,\mathbf{v}) \\ &\quad= U^\dagger(R)\begin{pmatrix} 1 & \boldsymbol{\sigma}\cdot\mathbf{v}/2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \tilde{E} + \frac{1}{2}m_0v^2 + \mathbf{v}\cdot R(\tilde{\mathbf{p}})\,, & -\boldsymbol{\sigma}\cdot[R(\tilde{\mathbf{p}}) + m_0\mathbf{v}] \\ -\boldsymbol{\sigma}\cdot[R(\tilde{\mathbf{p}}) + m_0\mathbf{v})]\,, & 2m_0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \boldsymbol{\sigma}\cdot\mathbf{v}/2 & 1 \end{pmatrix}U(R) \\ &\quad= U^\dagger(R)\begin{pmatrix} 1 & \boldsymbol{\sigma}\cdot\mathbf{v}/2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \tilde{E} + \mathbf{v}\cdot R(\tilde{\mathbf{p}}) - \frac{1}{2}\,(\boldsymbol{\sigma}\cdot R(\tilde{\mathbf{p}}))\,(\boldsymbol{\sigma}\cdot\mathbf{v}))\,, & -\boldsymbol{\sigma}\cdot R(\tilde{\mathbf{p}}) \\ -\boldsymbol{\sigma}\cdot R(\tilde{\mathbf{p}})\,, & 2m_0 \end{pmatrix}U(R) \\ &\quad= U^\dagger(R)\begin{pmatrix} \tilde{E} & -\boldsymbol{\sigma}\cdot R(\tilde{\mathbf{p}}) \\ -\boldsymbol{\sigma}\cdot R(\tilde{\mathbf{p}}) & 2m_0 \end{pmatrix}U(R) = \begin{pmatrix} \tilde{E} & -\boldsymbol{\sigma}\cdot\tilde{\mathbf{p}} \\ -\boldsymbol{\sigma}\cdot\tilde{\mathbf{p}} & 2m_0 \end{pmatrix} = \mathcal{D}^{(+)}(\mathbf{x},t)\,, \end{aligned}$$

which is what we were trying to prove. □

This means that the particle Pauli spinors transform according to:

$$\psi^{(+)\prime}(\mathbf{x}',t') = e^{if(\mathbf{x},t)/\hbar}\,\Lambda^{(+)}(R,\mathbf{v})\,\psi^{(+)}(\mathbf{x},t)\,, \tag{10.24}$$

and satisfy Pauli's equation in the transformed frame.

10.2.2 Spinor antiparticles

In non-relativistic theory, antiparticles are described as particles with negative values of the Galilean central charge $m = -m_0 < 0$. This negative value of m, however, is *not to be interpreted as negative mass*. Rather, we will have to put off the question of interpretation until later. For now, we consider it as a parameter in our theory. For $s = 1/2$, the Pauli equation for these negative m particles becomes:

$$\begin{pmatrix} -2m_0 & -\hbar\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}/i \\ -\hbar\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}/i & i\hbar\,\partial/\partial t \end{pmatrix}\begin{pmatrix} \chi^{(-)}(\mathbf{x},t) \\ \phi^{(-)}(\mathbf{x},t) \end{pmatrix} = 0\,, \tag{10.25}$$

where $\phi^{(-)}(\mathbf{x},t)$ and $\chi^{(-)}(\mathbf{x},t)$ are again two-component spinors. The $(-)$ superscript indicates that the phase m is negative. The solution of Eq. (10.25) is given by:

$$\chi^{(-)}(\mathbf{x},t) = -\frac{1}{2m_0}\,\frac{\hbar}{i}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,\phi^{(-)}(\mathbf{x},t)\,,\qquad i\hbar\,\frac{\partial}{\partial t}\,\phi^{(-)}(\mathbf{x},t) = \frac{\hbar}{i}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,\chi^{(-)}(\mathbf{x},t)\,, \tag{10.26}$$

which leads to the second order Schrödinger wave equation with $m < 0$ for the spinor $\phi^{(-)}(\mathbf{x},t)$:

$$\frac{\hbar^2}{2m_0}\,(\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,)^2\,\phi^{(-)}(\mathbf{x},t) = \frac{\hbar^2}{2m_0}\,I\,\nabla^2\,\phi^{(-)}(\mathbf{x},t) = i\hbar\,\frac{\partial}{\partial t}\,\phi^{(-)}(\mathbf{x},t)\,. \tag{10.27}$$

Solutions of the antiparticle wave equation also satisfy a conservation equation, given by:

$$\frac{\partial\rho^{(-)}(\mathbf{x},t)}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{j}^{(-)}(\mathbf{x},t) = 0\,, \tag{10.28}$$

with

$$\begin{aligned}\rho^{(-)}(\mathbf{x},t) &= \phi^{(-)\,\dagger}(\mathbf{x},t)\,\phi^{(-)}(\mathbf{x},t) = \psi^{(-)\,\dagger}(\mathbf{x},t)\,P^{(-)}\,\psi^{(-)}(\mathbf{x},t)\,, \\ \mathbf{j}^{(-)}(\mathbf{x},t) &= \phi^{(-)\,\dagger}(\mathbf{x},t)\,\boldsymbol{\sigma}\,\chi^{(-)}(\mathbf{x},t) + \chi^{(-)\,\dagger}(\mathbf{x},t)\,\boldsymbol{\sigma}\,\phi^{(-)}(\mathbf{x},t) \\ &= -\frac{\hbar}{2m_0\,i}\,\big[\,\phi^{(-)\,\dagger}(\mathbf{x},t)\,(\,\boldsymbol{\nabla}\,\phi^{(-)}(\mathbf{x},t)\,) - (\,\boldsymbol{\nabla}\,\phi^{(-)\,\dagger}(\mathbf{x},t)\,)\,\phi^{(-)}(\mathbf{x},t)\,\big] + \boldsymbol{\nabla}\times\mathbf{s}^{(-)}(\mathbf{x},t)\end{aligned} \tag{10.29}$$

where

$$P^{(-)} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\,,\qquad \mathbf{s}^{(-)}(\mathbf{x},t) = -\frac{\hbar}{2m_0}\,\big[\,\phi^{(-)\,\dagger}(\mathbf{x},t)\,\boldsymbol{\sigma}\,\phi^{(-)}(\mathbf{x},t)\,\big]\,, \tag{10.30}$$

is the spin probability density. These equations are consistent with identifying a negative charge to the electric current conservation equation.

Exercise 19. Establish Eqs. (10.28) and (10.29) from the Pauli equation for antiparticles.

Definition 25. Let us now define a matrix differential operator for the antiparticle equation. We write:

$$\mathcal{D}^{(-)}(\mathbf{x},t) = \begin{pmatrix} -2m_0 & -\boldsymbol{\sigma}\cdot\tilde{\mathbf{p}} \\ -\boldsymbol{\sigma}\cdot\tilde{\mathbf{p}} & \tilde{E} \end{pmatrix}\,,\qquad \psi^{(-)}(\mathbf{x},t) = \begin{pmatrix} \chi^{(-)}(\mathbf{x},t) \\ \phi^{(-)}(\mathbf{x},t) \end{pmatrix}\,, \tag{10.31}$$

where again $\tilde{E} = i\hbar\,\partial/\partial t$ and $\tilde{\mathbf{p}} = \hbar\,\boldsymbol{\nabla}/i$. Then Eq. (10.25) becomes:

$$\mathcal{D}^{(-)}(\mathbf{x},t)\,\psi^{(-)}(\mathbf{x},t) = 0\,. \tag{10.32}$$

For spinors, we define a charge conjugation matrix operator $C\mathcal{K}$ where $\mathcal{K}$ is the charge conjugate operator on functions and C is the matrix defined by:

$$C = C^\dagger = C^T = C^* = C^{-1} = \begin{pmatrix} 0 & i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix}\,. \tag{10.33}$$

This operator has the property of transforming the complex conjugate of the antiparticle Pauli equation into the particle Pauli equation:

$$C\,\mathcal{K}\,\mathcal{D}^{(-)}(\mathbf{x},t)\,\mathcal{K}^{-1}\,C^{-1} = -\mathcal{D}^{(+)}(\mathbf{x},t)\,. \tag{10.34}$$

Therefore:

$$C\,\mathcal{K}\,\mathcal{D}^{(-)}(\mathbf{x},t)\,\mathcal{K}^{-1}\,C^{-1}\,C\,\mathcal{K}\,\psi^{(-)}(\mathbf{x},t) = -\mathcal{D}^{(+)}(\mathbf{x},t)\,C\,\mathcal{K}\,\psi^{(-)}(\mathbf{x},t) = 0\,, \tag{10.35}$$

so that

$$\psi^{(+)}(\mathbf{x},t) = C\,\mathcal{K}\,\psi^{(-)}(\mathbf{x},t) = C\,\psi^{(-)\,*}(\mathbf{x},t)\,. \tag{10.36}$$

In component form, this means that complex conjugate solutions to the antiparticle Pauli equation, with negative values of m, can be interpreted as solutions of the particle Pauli equation with positive m and the upper and lower components reversed, that is:

$$\begin{aligned}\phi^{(+)}(\mathbf{x},t) &= i\sigma_2\,\phi^{(-)\,*}(\mathbf{x},t)\,,\\ \chi^{(+)}(\mathbf{x},t) &= -i\sigma_2\,\chi^{(-)\,*}(\mathbf{x},t)\,.\end{aligned} \tag{10.37}$$

Exercise 20. Using the solutions for $\chi^{(+)}(\mathbf{x},t)$ and $\chi^{(-)}(\mathbf{x},t)$ given in Eqs. (10.15) and (10.26), show that the last equation in (10.37) is consistent with the first equation.

Galilean transformation of solutions of the Pauli equation for negative m can be obtained by results for positive m. We first note that:

$$C\,\mathcal{K}\,\Lambda^{(\pm)}(R,\mathbf{v})\,\mathcal{K}^{-1}\,C^{-1} = \Lambda^{(\mp)}(R,\mathbf{v})\,. \tag{10.38}$$

Then from Theorem 29, it is easy to show that:

$$[\,\Lambda^{(-)}(R,\mathbf{v})\,]^{\dagger}\,e^{if(\mathbf{x},t)/\hbar}\,\mathcal{D}^{(-)}(\mathbf{x}',t')\,e^{-if(\mathbf{x},t)/\hbar}\,\Lambda^{(-)}(R,\mathbf{v}) = \mathcal{D}^{(-)}(\mathbf{x},t)\,, \tag{10.39}$$

where $f(\mathbf{x},t)$ is given as before in Eq. (10.9).

Exercise 21. Prove Eq. (10.39).

So the antiparticle Pauli spinors transform according to:

$$\psi^{(-)\,\prime}(\mathbf{x}',t') = e^{-if(\mathbf{x},t)/\hbar}\,\Lambda^{(-)}(R,\mathbf{v})\,\psi^{(-)}(\mathbf{x},t)\,, \tag{10.40}$$

and satisfy the antiparticle Pauli's equation in the transformed frame.

Since solutions to the particle and antiparticle equations transform differently, the best we can do is to define an eight component spinor with each component the four component particle and antiparticle solutions, as we did for scalars:

$$\Psi(\mathbf{x},t) = \begin{pmatrix}\psi^{(+)}(\mathbf{x},t)\\ \psi^{(-)}(\mathbf{x},t)\end{pmatrix}\,, \tag{10.41}$$

which transform under Galilean transformations as:

$$\Psi'(\mathbf{x}',t') = T(\mathbf{x},t)\,\Psi(\mathbf{x},t)\,, \tag{10.42}$$

where

$$T(\mathbf{x},t) = \begin{pmatrix}e^{+if(\mathbf{x},t)/\hbar}\,\Lambda^{(+)}(R,\mathbf{v}) & 0\\ 0 & e^{-if(\mathbf{x},t)/\hbar}\,\Lambda^{(-)}(R,\mathbf{v})\end{pmatrix}\,. \tag{10.43}$$

10.3 Vectors

We construct wave equations for particles of spin one by the method suggested by Dirac and developed by Bargmann and Wigner for relativisitc particles of any spin. We discuss that method for the Poincaré group in Appendix **??**. Here, we need to assure that the equation will be invaraint under Galiliean transformations rather than Lorentz transformation. Following Dirac's method, we propose a matrix-spinor non-relativistic wave function $\Psi_{\alpha_1,\alpha_2}(x)$ for positive mass particles which satisfies the equation:

$$\begin{aligned} \mathcal{D}_{\alpha_1,\alpha_1'}(\mathbf{x},t)\,\Psi_{\alpha_1',\alpha_2}(x) &= 0\,, \\ \mathcal{D}_{\alpha_2,\alpha_2'}(\mathbf{x},t)\,\Psi_{\alpha_1,\alpha_2'}(x) &= 0\,. \end{aligned} \tag{10.44}$$

where $\mathcal{D}_{\alpha,\alpha'}(\mathbf{x},t)$ is given in Eq. (10.20). Here we have dropped the (+) designation for positive mass solutions.

We will work this out the same way we did for the Proca equation in Appendix **??**.

10.4 Massless wave equations

10.4.1 Massless scalers

This must be an equation of the form:

$$\nabla^2\phi(\mathbf{x}) = 0\,, \tag{10.45}$$

which has the solution:

$$\phi(\mathbf{x}) = \frac{1}{r}\,. \tag{10.46}$$

A candidate for the realization of this must be a scalar graviton. This must be Newton's theory of gravity, with an instantaneous interaction?

10.4.2 Massless vectors

Well, surely this is electrodynamics with an infinite velocity of light. It is possible to work this out from the massive vector field of Section 10.3.

Quite a bit to do here yet!

References

Chapter 11

Supersymmetry

We have seen examples of specific non-observable qualities of Nature. The essential two-valuedness of the non-relativistic electron, described by spin, is one such example. In fact, any two-level quantum system can be described by essentially non-observable variables. We have learned to describe these systems by **Grassmann** variables. In some systems, a symmetry can exist between the Grassmannian variables and the ordinary variables. We discuss in this chapter Grassmann variables and supersymmetry transformations.

11.1 Grassmann variables

Grassmann variables are *classical* variables which obey an anti-commuting algebra. In this respect they share things in common with Fermi anti-commuting *operators*, but are considered to be the classical variable which is mapped to a quantum operator in much the same way that the classical coordinate q is mapped to a quantum operator Q. Grassmann variables have unusual properties, some of which are discussed here.

Definition 26 (Grassmann variables)**.** A set of N quantities θ_i, $i = 1, 2, \ldots, N$ are Grassmann variables if they obey the anti-commutation relations:

$$\{\, \theta_i, \theta_j \,\} = 0 \,. \tag{11.1}$$

Grassmann variables *commute* with all other classical variables.

This definition implies that $\theta_i^2 = 0$ for all i. Functions of Grassmann variables are defined by their power series expansions. For example, any function $f(\theta)$ of a single Grassmann variable which has a Taylor series expansion about the origin can be written as:

$$f(\theta) = f(0) + f'(0)\,\theta \,, \tag{11.2}$$

since $\theta^2 = 0$. Functions of two or more Grassmann variables get more complicated. For example for two Grassmann variables, $f(\theta_1, \theta_2)$ is of the form:

$$f(\theta_1, \theta_2) = f(0,0) + f_1(0,0)\,\theta_1 + f_2(0,0)\,\theta_2 + f_{1,2}(0,0)\,\theta_1\theta_2 \,. \tag{11.3}$$

We also define derivatives of Grassmann variables in the following definition.

Definition 27 (differentiation)**.** Derivatives of Grassmann variables are taken to be *left*-acting and anti-commute.

$$\{\, \partial_i, \theta_j \,\} = \delta_{ij} \,, \qquad \{\, \partial_i, \partial_j \,\} = 0 \,, \qquad \partial_i \equiv \frac{\partial}{\partial \theta_i} \,. \tag{11.4}$$

This means, for example, that:

$$\partial_i(\, \theta_j\theta_k \,) = \delta_{ij}\,\theta_k - \delta_{ik}\,\theta_j \,. \tag{11.5}$$

Integration of Grassmann variables has some unusual properties, which are given in the next two definitions.

Definition 28 (integration)**.** The differential obeys the following rules:

$$\{ \mathrm{d}\theta_i, \theta_j \} = 0 \,, \qquad \{ \mathrm{d}\theta_i, \mathrm{d}\theta_j \} = 0 \,, \qquad \{ \mathrm{d}\theta_i, \partial_j \} = \delta_{ij} \,. \tag{11.6}$$

Integrals are defined by:

$$\int \mathrm{d}\theta = 0 \,, \qquad \int \mathrm{d}\theta \, \theta = 1 \,. \tag{11.7}$$

The integration rules mean that a Grassmann Dirac δ-function can be defined by:

$$\delta(\theta) = \theta \,. \tag{11.8}$$

Up until now, we have been considering real Grassmann variables. If θ_1 and θ_2 are two real Grassmann variables, complex Grassmann variables can be defined as follows:[1]

$$\theta = (\theta_1 + i\theta_2)/\sqrt{2} \,, \qquad \theta_1 = (\theta + \theta^*)/\sqrt{2} \,, \tag{11.9}$$

$$\theta^* = (\theta_1 - i\theta_2)/\sqrt{2} \,, \qquad \theta_1 = (\theta - \theta^*)/i\sqrt{2} \,. \tag{11.10}$$

For the derivatives, we have:

$$\partial_\theta = (\partial_1 - i\partial_2)/\sqrt{2} \,, \qquad \partial_1 = (\partial_\theta + \partial^*_\theta)/\sqrt{2} \,, \tag{11.11}$$

$$\partial^*_\theta = (\partial_1 + i\partial_2)/\sqrt{2} \,, \qquad \partial_2 = i(\partial_\theta + \partial^*_\theta)/\sqrt{2} \,. \tag{11.12}$$

The complex variables satisfy the algebra:

$$\{ \theta, \theta \} = \{ \theta^*, \theta^* \} = \{ \theta, \theta^* \} = 0 \,. \tag{11.13}$$

Integrals over complex Grassmann variables are given by:

$$\int \mathrm{d}\theta = \int \mathrm{d}\theta^* = 1 \,, \qquad \int \mathrm{d}\theta \, \theta = \int \mathrm{d}\theta^* \, \theta^* = 0 \,, \tag{11.14}$$

and

$$\mathrm{d}^2\theta = \mathrm{d}\theta \, \mathrm{d}\theta^* = i \, \mathrm{d}\theta_1 \, \mathrm{d}\theta_2 \,. \tag{11.15}$$

The complex conjugate of two Grassmann variables is defined by:

$$[\theta_1 \theta_2]^* = \theta^*_2 \, \theta^*_1 \,, \tag{11.16}$$

which is similar to the Hermitian adjoint operation for matrices.

11.2 Superspace and the 1D-N supersymmetry group

In this section, we discuss supersymmetry in one dimension with N real Grassmann variables. Super-space consists of

$$s = (t, \theta_r) \,, \tag{11.17}$$

with $r = 1, \ldots, N$. Here t is a real variable and θ_r are real Grassmann variables which anticommute: $\{ \theta_r, \theta_{r'} \} = \delta_{rr'}$. A supersymmetry transformation is given by:

$$\begin{aligned} t' &= t + \tau + i \, \chi_r \theta_r \,, \\ \theta'_r &= \theta_r + \chi_r \,. \end{aligned} \tag{11.18}$$

[1] We use a factor of $1/\sqrt{2}$ for convience here.

This represents a displacement (χ_r) in Grassmannian space combined with a shift in clocks (τ) and an additional shift in clocks porportional to the product of the Grassmann shift and the position in Grassman space.

We demand that no experiment can be performed on the system which can detect the difference between these two coordinate systems. That is, this is a symmetry of Nature.

We start by proving that the elements $g = (\tau, \chi_r)$ form a group. In order to do this, we need to find the group multiplication rule: $g'' = g'g$, and find the identity and inverse element. We first establish the group composition rule:

Theorem 30 (Composition rule)**.** *The composition rule for 1D-N supersymmetry is:*

$$\begin{aligned} \tau'' &= \tau' + \tau + i\,\chi'_r\chi_r\,, \\ \chi''_r &= \chi'_r + \chi_r\,. \end{aligned} \tag{11.19}$$

Proof. We first note that

$$\theta''_r = \theta'_r + \chi'_r = \theta_r + \chi_r + \chi'_r \equiv \theta_r + \chi''_r\,,$$

where $\chi''_r = \chi'_r + \chi_r$, which establishes the χ_r composition rule. Next, we find:

$$\begin{aligned} t'' &= t' + \tau' + i\,\chi'_r\theta'_r \\ &= t + \tau + i\,\chi_r\theta_r + \tau' + i\,\chi'_r(\,\theta_r + \chi_r\,) \\ &= t + \tau' + \tau + i\,\chi'_r\chi_r + i\,\chi''_r\theta_r \\ &\equiv t + \tau'' + i\,\chi''_r\theta_r\,, \end{aligned}$$

where $\tau'' = \tau' + \tau + i\,\chi'_r\chi_r$. This completes the proof. □

We further note that the unit element $1 = (0, 0)$ does nothing to the transformation, and that $g^{-1} = (-\tau, -\chi_r)$, because of the Grassmann nature of the χ_r variables. So the elements $g = (\tau, \chi_r)$ of the 1D-N supersymmetry transformation form a group.

11.3 1D-N supersymmetry transformations in quantum mechanics

Recall that the state of a quantum system is described by a ray in Hilbert space. Two vectors $|\,\Psi\,\rangle$ and $|\,\Psi'\,\rangle$ in Hilbert space belong to the same ray if they differ by a phase, $|\,\Psi'\,\rangle = e^{i\phi}|\,\Psi\,\rangle$. Symmetry transformations are represented in quantum mechanics by unitary or anti-unitary operators acting on rays. So in this section, we want to find unitary transformations that represent supersymmetry transformations in ordinary space. In technical terms, we want to find representations for the unitary covering group for the supersymmetry group.

Let $U(g)$ be the unitary transformation which takes a vector $|\,\Psi\,\rangle$ in the ray $\mathcal{R}$ to a vector $|\,\Psi(g)\,\rangle$ in the ray $\mathcal{R}(g)$ as:

$$|\,\Psi(g)\,\rangle = U(g)\,|\,\Psi\,\rangle\,, \tag{11.20}$$

Any vector in the same ray $\mathcal{R}(g)$ describes the same physical system in the transformed system g. There is one special state, called the "vacuum" state or ground state of the system, which is invariant under supersymmetry transformations. This means that:

$$U(g)\,|\,0\,\rangle = |\,0\,\rangle\,. \tag{11.21}$$

We will use this fact later.

The product of two supersymmetry transformations, $\mathcal{R} \to \mathcal{R}(g) \to \mathcal{R}(g'g)$ gives a vector in the ray $\mathcal{R}(g'g)$,

$$|\,\Psi(g'g)\,\rangle = U(g')\,|\,\Psi(g)\,\rangle = U(g')\,U(g)\,|\,\Psi\,\rangle\,.$$

However the direct transformation from the ray $\mathcal{R} \to \mathcal{R}(g'g)$ gives:

$$|\,\Psi'(g'g)\,\rangle = U(g')\,|\,\Psi\,\rangle\,.$$

But $|\,\Psi(g'g)\,\rangle$ and $|\,\Psi'(g'g)\,\rangle$ have to be in the same ray since they describe the same physical system, so $|\,\Psi(g'g)\,\rangle = e^{i\phi(g',g)}|\,\Psi'(g'g)\,\rangle$. Therefore the group multiplication rule for unitary operators representing supersymmetry transformations in Hilbert space is given by:

$$U(g')\,U(g) = e^{i\phi(g',g)}\,U(g'g)\,. \tag{11.22}$$

Representations of operators which obey (11.22) are called *projective* representations. The supersymmetry group is a *continuous* projective group of infinite dimension.

The unit element is $U(1) = 1$. So using the group composition rule (11.22), unitarity requires that:

$$U^\dagger(g)U(g) = U^{-1}(g)U(g) = U(g^{-1})U(g) = e^{i\phi(g^{-1},g)}\,U(1,0) = 1\,. \tag{11.23}$$

provided that $\phi(g^{-1},g) = 0$. The associative law for group transformations,

$$U(g'')\,(U(g')U(g)) = (U(g'')U(g'))\,U(g)\,,$$

requires that the phases satisfy:

$$\phi(g'',g'g) + \phi(g',g) = \phi(g'',g') + \phi(g''g',g)\,, \tag{11.24}$$

with $\phi(1,1) = \phi(1,g) = \phi(g,1) = \phi(g^{-1},g) = 0$. Note that the phase rule (11.24) can be satisfied by any $\phi(g',g)$ of the form

$$\phi(g',g) = \alpha(g'g) - \alpha(g') - \alpha(g)\,. \tag{11.25}$$

Then the phase can be eliminated by a trivial change of phase of the unitary transformation, $\bar{U}(g) = e^{i\alpha(g)}U(g)$. Thus two phases $\phi(g',g)$ and $\phi'(g',g)$ which differ from each other by functions of the form (11.25) are equivalent. Finding nontrival phases means that there are central charges in the algebra of the group.

For the 1D-N supersymmetry group, the phase is given by the following theorem:

Theorem 31 (1D-N supersymmetry phase)**.** *The phase is given by:*

$$\phi(g',g) = i\,\chi'_r\,M_{rr'}\,\chi_{r'}\,, \tag{11.26}$$

where $M_{rr'}$ is a real traceless $N \times N$ symmetric matrix.

Proof. Following a method due to Bargmann[1], we first note that the transformation rule is *linear* in χ_r. So it is obvious that $\phi(g'g)$ must be bilinear in χ_r and $\chi'_{r'}$. So we make the ansatz:

$$\phi(g',g) = i\,\chi'_r\,M_{rr'}\,\chi_{r'}\,, \tag{11.27}$$

where $M_{rr'}$ is a general $N \times N$ matrix. So we find:

$$\begin{aligned}
\phi(g'',g'g) &= i\,\chi''_r\,M_{rr'}\,(\,\chi'_{r'} + \chi_{r'}\,)\,,\\
\phi(g',g) &= i\,\chi'_r\,M_{rr'}\,\chi_{r'}\,,\\
\phi(g'',g') &= i\,\chi''_r\,M_{rr'}\,\chi'_{r'}\,,\\
\phi(g''g',g) &= i\,(\,\chi''_r + \chi'_r\,)\,M_{rr'}\,\chi_{r'}\,,
\end{aligned}$$

from which we see that the phase rule, Eq. (11.24), is satisfied. We also note that due to the properties of Grassmann varibles,

$$\phi^*(g',g) = i\,\chi'_r\,M^*_{rr'}\,\chi_{r'}\,, \tag{11.28}$$

so in order for the phase $\phi(g',g)$ to be real, $M^*_{rr'} = M_{rr'}$ must be real. Next, we write $M_{rr'}$ as a sum of symmetric and antisymmetric matrices:

$$\begin{aligned} M_{rr'} &= \frac{1}{2}\,[\,M_{rr'} + M_{r'r}\,] - \frac{1}{N}\,\delta_{r,r'}\,\mathrm{Tr}[\,M\,] + \frac{1}{2}\,[\,M_{rr'} - M_{r'r}\,] + \frac{1}{N}\,\delta_{r,r'}\,\mathrm{Tr}[\,M\,] \\ &= M^{ST}_{rr'} + M^{A}_{rr'} + \frac{1}{N}\,\delta_{r,r'}\,\mathrm{Tr}[\,M\,]\,. \end{aligned}$$

where $M^{ST}_{rr'}$ is the traceless symmetric part of M and $M^{A}_{rr'}$ is the antisymmetric part of M. Now the antisymmetric part is a trival phase, because if we set

$$\alpha_1(g) = \frac{i}{2}\,\chi_r\,M_{rr'}\,\chi_{r'}\,, \tag{11.29}$$

Then using the composition rule,

$$\begin{aligned} \alpha_1(g'') &= \frac{i}{2}\,\chi''_r\,M_{rr'}\,\chi''_{r'} \\ &= \frac{i}{2}\,(\chi'_r + \chi_r)\,M_{rr'}\,(\chi'_{r'} + \chi_{r'}) \\ &= \alpha_1(g') + \alpha_1(g) + \frac{i}{2}\,\chi'_r\,M_{rr'}\,\chi_{r'} + \frac{i}{2}\,\chi_r\,M_{rr'}\,\chi'_{r'} \\ &= \alpha_1(g') + \alpha_1(g) + \frac{i}{2}\,\chi'_r\,[\,M_{rr'} - M_{r'r}\,]\,\chi_{r'} \end{aligned} \tag{11.30}$$

So

$$i\,\chi'_r\,M^{A}_{rr'}\,\chi_{r'} = \alpha_1(g'') - \alpha_1(g') - \alpha_1(g)\,, \tag{11.31}$$

and is thus a trivial phase and can be removed. For the trace part, we set:

$$\alpha_2(g) = \frac{\mathrm{Tr}[\,M\,]}{N}\,\tau\,. \tag{11.32}$$

Then from the composition rule for τ in Eq. (11.19), we find:

$$i\,\frac{\mathrm{Tr}[\,M\,]}{N}\chi'_r\chi_r = \alpha_2(g'') - \alpha_2(g') - \alpha_2(g)\,. \tag{11.33}$$

So the trace part is also a trival phase and can be removed. This leaves only the symmetric traceless part,

$$\phi(g',g) = i\,\chi'_r\,M^{ST}_{rr'}\,\chi_{r'}\,,$$

which is what we were trying to prove. From now on, we drop the "ST" labeling on $M_{rr'}$, and just keep in mind that $M_{rr'}$ is an $N\times N$ traceless symmetric matrix with $N(N+1)/2-1$ independent real numbers which commute with all generators of the group. □

Remark 23. We note that:

$$\phi(1,g) = \phi(g,1) = 0\,, \tag{11.34}$$

and, using the fact that $M_{rr'}$ is traceless and symmetric, we find:

$$\phi(g^{-1},g) = -i\,\chi_r\,M_{rr'}\,\chi_{r'} = +i\,\chi_{r'}\,M_{rr'}\,\chi_r = +i\,\chi_{r'}\,M_{r'r}\,\chi_r = -\phi(g^{-1},g)\,, \tag{11.35}$$

so that $\phi(g^{-1},g) = \phi(g,g^{-1}) = 0$. We will use these relations below.

11.4 Supersymmetric generators

For infinitesimal transformations, the generators H and Q_r of supersymmetry transformations are defined by:

$$U(1+\Delta g) = 1 + i\,\Delta\tau\, H + \Delta\chi_r\, Q_r + \cdots \tag{11.36}$$

Here the generator of time displacements H is called the **hamiltonian** and the generators of Grassmann coordinate displacements Q_r are called **supercharges**. The supercharges Q_r *anti*commute with the Grassmann displacements $\chi_{r'}$,

$$\{\, Q_r, \chi_{r'} \,\} = 0\,, \tag{11.37}$$

but *not necessarily with themselves.* Since, from Eq. (11.21), the vacuum state is invariant under supersymmetry transformations, we must have:

$$H\,|\,0\,\rangle = 0\,, \qquad Q_r\,|\,0\,\rangle = 0\,, \qquad \text{for all } r = 1,\ldots,N. \tag{11.38}$$

Next, we work out the transformation properties of the group generators. We do this in the following theorem:

Theorem 32 (Group transformations)**.** *The group generators transform according to the rules:*

$$\begin{aligned} U^{-1}(g)\, H\, U(g) &= H\,, \\ U^{-1}(g)\, Q_r\, U(g) &= Q_r - 2\,(\,H\,\delta_{rr'} + M_{rr'}\,)\,\chi_{r'}\,. \end{aligned} \tag{11.39}$$

Proof. We start by considering the transformation:

$$U(g^{-1})\,U(1+\Delta g')\,U(g) = e^{i\beta(g,\Delta g')}\,U(1+\Delta g'')\,, \tag{11.40}$$

where $\Delta g'' = g^{-1}\,\Delta g'\,g$, and where the phase $\beta(g,\Delta g')$ is given by:

$$\beta(g,\Delta g') = \phi(\,g^{-1},(1+\Delta g')\,g\,) + \phi(\,1+\Delta g', g\,) \tag{11.41}$$

We can simplify this expression using the phase rule Eq. (11.24), with the substitutions:

$$\begin{aligned} g'' &\mapsto g \\ g' &\mapsto g^{-1} \\ g &\mapsto (\,1+\Delta g'\,)\,g\,, \end{aligned} \tag{11.42}$$

so that:

$$\phi(\,g, g^{-1}(1+\Delta g')g\,) + \phi(\,g^{-1},(1+\Delta g')g\,) = \phi(\,g,g^{-1}\,) + \phi(\,gg^{-1},(1+\Delta g')g\,)\,, \tag{11.43}$$

But using the results in remark 23, we find:

$$\phi(\,g,g^{-1}\,) = \phi(\,1,(1+\Delta g')g\,) = 0\,. \tag{11.44}$$

Then (11.43) becomes:

$$\phi(\,g^{-1},(1+\Delta g')g\,) = -\phi(\,g,g^{-1}(1+\Delta g')g\,) = -\phi(\,g,1+\Delta g''\,)\,. \tag{11.45}$$

So the phase $\beta(g,\Delta g')$ in Eq. (11.41) can be written as:

$$\beta(g,\Delta g') = \phi(\,1+\Delta g', g\,) - \phi(\,g,1+\Delta g''\,)\,. \tag{11.46}$$

Now we need to work out the transformation $\Delta g'' = g^{-1}\,\Delta g'\,g := (\,\Delta\tau'', \Delta\chi_r''\,)$. 1For our case,

$$\begin{aligned} g &= (\tau,\chi_r) \\ \delta g &= (\Delta\tau, \Delta\chi_r) \\ g^{-1} &= (-\tau,-\chi_r)\,, \end{aligned} \tag{11.47}$$

So after some work, we find:

$$\begin{aligned}\Delta\tau'' &= \Delta\tau' + 2i\,\Delta\chi'_r\chi_r\,,\\ \Delta\chi''_r &= \Delta\chi'_r\,.\end{aligned}\tag{11.48}$$

So for the phase, we find:

$$\begin{aligned}\phi(\,1+\Delta g', g\,) &= i\,\Delta\chi'_r\,M_{rr'}\,\chi_{r'}\,,\\ \phi(\,g, 1+\Delta g''\,) &= i\,\chi_r\,M_{rr'}\,\Delta\chi''_r\\ &= i\,\chi_r\,M_{rr'}\,\Delta\chi'_r = -i\,\Delta\chi'_r\,M_{rr'}\,\chi_{r'}\,,\end{aligned}\tag{11.49}$$

since $M_{rr'}$ is *symmetric*. So the phase $\beta(g,\Delta g')$ becomes:

$$\beta(g,\Delta g') = 2i\,\Delta\chi'_r\,M_{rr'}\,\chi_{r'}\,.\tag{11.50}$$

Now using (11.36), we find:

$$\begin{aligned}U(1+\Delta g') &= 1 + i\,\Delta\tau'\,H + \Delta\chi'_r\,Q_r + \cdots\\ U(1+\Delta g'') &= 1 + i\,\Delta\tau''\,H + \Delta\chi''_r\,Q_r + \cdots\end{aligned}\tag{11.51}$$

Putting all this into Eq. (11.40), and expanding the phase out to first order gives:

$$\begin{aligned}1 + i\Delta\tau'\,U^{-1}(g)\,H\,U(g) + \Delta\chi'_r\,U^{-1}(g)\,Q_r\,U(g) + \cdots&\\ = 1 + i\Delta\tau''\,H\ + \Delta\chi''_r\,Q_r - 2\,\Delta\chi'_r\,M_{rr'}\,\chi_{r'} + \cdots&\\ = 1 + i\Delta\tau'\,H\ + \Delta\chi'_r\,\big[\,Q_r - 2\,(\,H\,\delta_{rr'} + M_{rr'}\,)\,\chi_{r'}\,\big] + \cdots&\end{aligned}\tag{11.52}$$

comparing coefficients of $\Delta\tau'$ and $\Delta\chi'_r$ gives:

$$\begin{aligned}U^{-1}(g)\,H\,U(g) &= H\,,\\ U^{-1}(g)\,Q_r\,U(g) &= Q_r - 2\,(\,H\,\delta_{rr'} + M_{rr'}\,)\,\chi_{r'}\,,\end{aligned}\tag{11.53}$$

which is the result we were trying to prove. □

We can now find the algebra obeyed by the generators from the results of Theorem 32. This algebra is stated in the following theorem:

Theorem 33 (Group algebra)**.** *The group generators transform according to the rules:*

$$\begin{aligned}[\,H, Q_r\,] &= 0\,,\\ \{\,Q_r, Q_{r'}\,\} &= 2\,(\,H\,\delta_{rr'} + M_{rr'}\,)\,.\end{aligned}\tag{11.54}$$

The $N(N+1)/2-1$ values of the traceless symmetric matrix $M_{rr'}$ are called the ***central charges*** *of the algebra.*

Proof. We set $g = 1+\Delta g$ in Eq. (11.39), and compare both sides of the equations. For the first equation, we find:

$$\big(\,1 - i\,\Delta\tau\,H - \Delta\chi_{r'}\,Q_{r'} + \cdots\big)\,H\,\big(\,1 + i\,\Delta\tau\,H + \Delta\chi_{r'}\,Q_{r'} + \cdots\big) = H\,,$$

So this means that $[\,H, Q_r\,] = 0$. In the second equation, we find:

$$\begin{aligned}\big(\,1 - i\,\Delta\tau\,H - \Delta\chi_{r'}\,Q_{r'} + \cdots\big)\,Q_r\,\big(\,1 + i\,\Delta\tau\,H + \Delta\chi_{r'}\,Q_{r'} + \cdots\big)&\\ = Q_r - 2\,(\,H\,\delta_{rr'} + M_{rr'}\,)\,\Delta\chi_{r'}\,,&\end{aligned}$$

So in addition to the first result, we find here that:

$$-\left[\, Q_{r'}\, Q_r + Q_r\, Q_{r'} \,\right] \Delta\chi_{r'} = -2\,(\, H\,\delta_{rr'} + M_{rr'}\,)\,\Delta\chi_{r'}\ .$$

which gives the anticommutator:

$$\{\, Q_r, Q_{r'}\,\} = 2\,(\, H\,\delta_{rr'} + M_{rr'}\,)\ ,$$

which completes the proof. □

Remark 24. Note that from the second of Eq. (11.54), we find $\{\, Q_r, Q_{r'}\,\} = 2\,M_{rr'}$ for $r \neq r'$, and $Q_r^2 = H + M_{rr}$ for $r = r'$.

Example 30 (N=1)**.** For $N = 1$, there are no central charges and only one supercharge Q, obeying the algebra:

$$\{\, Q, Q\,\} = 2\,H\ , \tag{11.55}$$

which means that H can be factored by the (real) supercharge operator: $H = Q^2$. We will see some specific illustrations of $N = 1$ supersymmetry models later on.

Example 31 (N=2)**.** For $N = 2$, we set the central charge matrix to:

$$M = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}\ , \tag{11.56}$$

where a and b are real. Now it is useful to define *complex* supercharges by:

$$\begin{aligned} \tilde{Q}_1 = \tilde{Q}\ &= \frac{1}{\sqrt{2}}\,(\, Q_1 + i\,Q_2\,)\ , \\ \tilde{Q}_2 = \tilde{Q}^*&= \frac{1}{\sqrt{2}}\,(\, Q_1 - i\,Q_2\,)\ , \end{aligned} \tag{11.57}$$

so that we can write:

$$\tilde{Q}_s = U_{sr}\,Q_r\ , \qquad \tilde{Q}^*_{s'} = U^*_{s'r'}\,Q_{r'} = Q_{r'}\,[U^\dagger]_{r's'}\ . \tag{11.58}$$

where the unitary matrix U is defined by:

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}\ , \qquad U^\dagger = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}\ . \tag{11.59}$$

So we find:

$$\begin{aligned} \{\, \tilde{Q}_s, \tilde{Q}^*_{s'}\,\} &= U_{sr}\,\{\, Q_r, Q_{r'}\,\}\,U_{r's'} = 2\,U_{sr}\,(\, H\,\delta_{rr'} + M_{rr'}\,)\,[U^\dagger]_{r's'}\ , \\ &= 2\begin{pmatrix} H & z \\ z^* & H \end{pmatrix}\ , \end{aligned} \tag{11.60}$$

where $z = a + ib$. In this form, we have the striking result:

$$H = \frac{1}{2}\,(\, \tilde{Q}^*\,\tilde{Q} + \tilde{Q}\,\tilde{Q}^*\,)\ , \qquad \tilde{Q}^2 = z\ , \qquad \tilde{Q}^{*\,2} = z^*\ . \tag{11.61}$$

Of course, we also have:

$$[\, H, \tilde{Q}\,] = [\, H, \tilde{Q}^*\,] = 0\ , \tag{11.62}$$

and the central charges z and z^* commute with everything. In this form, we see that z is the (complex) normalization of the supercharge operator $\tilde{Q}$, and the hamiltonian has been *factored* by the supercharge and its complex conjugate so that it's eigenvalues must be non-negative. That is, if we write:

$$H\,|\,E\,\rangle = E\,|\,E\,\rangle\ , \tag{11.63}$$

$$\begin{array}{ccc} s=(t,\theta) & \xrightarrow{g=(\tau,\chi)} & s'=(t',\theta') \\ {\scriptstyle R}\downarrow & & \uparrow{\scriptstyle R^T} \\ \tilde{s}=(t,\tilde{\theta}) & \xrightarrow{\tilde{g}=(\tau,\tilde{\chi}')} & \tilde{s}'=(t',\tilde{\theta}') \end{array}$$

Figure 11.1: R-symmetry.

we see that:

$$E = \langle\, E \,|\, H \,|\, E \,\rangle = \frac{1}{2}\left[\, ||\, Q \,|\, E \,\rangle\,||^2 + ||\, Q^* \,|\, E \,\rangle\,||^2 \,\right] \geq 0\,. \tag{11.64}$$

For the vacuum state, we find

$$H \,|\, 0 \,\rangle = Q \,|\, 0 \,\rangle = Q^* \,|\, 0 \,\rangle = 0\,. \tag{11.65}$$

So for this state, $E = 0$. Moreover we note that since $[\, Q, H \,] = [\, Q^*, H \,] = 0$, the states $|\, E \,\rangle$, $Q \,|\, E \,\rangle$, and $Q^* \,|\, E \,\rangle$ all have the same energy E, and we say that they belong to the same **multiplet**.

11.5 R-symmetry

For $N > 1$, we see from the supersymmetry transformation given in Eq. (11.18) that the combination: $i\,\chi_r\theta_r$ can be considered to be an inner product of two real N-dimensional vectors composed of Grassmann variables. This inner product is invariant under simultaneous orthogonal transformations (rotations) of the coordinate system and the transformation parameters. This invariance is called **R-symmetry**. We explain this symmetry here. Let $\tilde{\theta}_r$, $r = 1, \ldots, N$ be a Grassmann coordinate system obtained from the original one by an orthogonal transformation:

$$\tilde{\theta}_r = R_{rr'}\,\theta_{r'}\,, \tag{11.66}$$

where R is real orthogonal matrix, $R\,R^T = 1$. Then a supersymmetry transfomation from a coordinate set $s = (t, \theta_r)$ to the coordinate set $s' = (t', \theta'_r)$, as described by the group parameters $g = (\tau, \chi_r)$, is the same as the transformation from a set $\tilde{s} = (t, \theta_r)$ to the set $\tilde{s}' = (t', \tilde{\theta}'_r)$ as described by the group parameters $\tilde{g} = (\tau, \tilde{\chi}_r)$, where:

$$\tilde{\chi}_r = R_{rr'}\,\chi_{r'}\,. \tag{11.67}$$

R-symmetry is illustrated in Fig. 11.1. Now since the central charge matrix $M_{rr'}$ is real and symmetric, it can be diagonalized by an orthogonal matrix. So given a set of central charges $M_{rr'}$ we can always bring it to diagonal form by an R transformation. That is, let R be such that

$$\chi_r\,M_{rr'}\,\chi_{r'} = \tilde{\chi}_s\,R_{rs}\,M_{rr'}\,R_{r',s'}\,\tilde{\chi}_{s'} = M_s\,\tilde{\chi}_s\,\tilde{\chi}_s\,, \tag{11.68}$$

where M_s are the eigenvalues of the central charge matrix,

$$R_{rs}\,M_{rr'}\,R_{r',s'} = M_s\,\delta_{s,s'}\,. \tag{11.69}$$

Since the trace is invariant under orthogonal transformations, the sum of the eigenvalues is zero: $\sum_s M_s = 0$. Defining

$$\tilde{Q}_s = R_{sr}\,Q_r\,. \tag{11.70}$$

In fact, for $N = 2$, we find that $M_s = \pm\sqrt{a^2 + b^2} = \pm|z|$. In the basis set defined by $\tilde{Q}_r$ and $\tilde{\chi}_r$ gives the group transformations:

$$\begin{aligned} U^{-1}(g)\,H\,U(g) &= H\,, \\ U^{-1}(g)\,\tilde{Q}_s\,U(g) &= \tilde{Q}_s - 2\,(\,H + M_s\,)\,\tilde{\chi}_s\,. \end{aligned} \tag{11.71}$$

and the algebra:

$$\begin{aligned} [\, H, \tilde{Q}_s \,] &= 0 \,, \\ \{\, \tilde{Q}_s, \tilde{Q}_{s'} \,\} &= 2\,(\, H + M_s \,)\,\delta_{ss'} \,. \end{aligned} \tag{11.72}$$

In this system, the phase is diagonal:

$$\phi(g', g) = i\,\chi_r\, M_{rr'}\,\chi_{r'} = i\, M_s\,\tilde{\chi}_s\,\tilde{\chi}_s \,. \tag{11.73}$$

So, in addition to the $N(N-1)/2$ components of an R-transformation which leave supersymmetry transformations invariant, the central charge matrix can always be brought to diagonal form with $N-1$ eigenvalues.

11.6 Extension of the supersymmetry group

In this section, we assume that we are working in a system of Grassmann coordinates in which the central charge matrix $M_{rr'}$ is diagonal. Then the phase factor (11.73) is linear in M_s, we can extend the group by promoting these quantities to be additional generators of the group and operators in Hilbert space. We introduce new (real) group parameters μ_s, with $s = 1, \ldots, N$, which transform according to the rule:

$$\mu''_{\ s} = \mu'_{\ s} + \mu_s + i\,\tilde{\chi}_s\,\tilde{\chi}_s \,, \qquad \text{(no sum over } s.) \tag{11.74}$$

Then we note that:

$$M_s\,\mu''_s - M_s\,\mu'_s - M_s\,\mu_s = i\,M_s\tilde{\chi}'_s\,\tilde{\chi}_s = \phi(g', g) \,. \tag{11.75}$$

(Here there is an implied sum over s.) So that the phase relation can be acheved by redefinition of the unitary transformation to be:

$$U_{\rm ext}(\tilde{g}) = U(g)\, e^{\,i\,M_s\,\mu_s} \tag{11.76}$$

where the extended group, $\mathcal{SUSY}_{\rm ext}$ now consists of the $1{+}2N$ elements: $\tilde{g} = (\,\tau, \tilde{\chi}_s, \mu_s\,)$, with the composition rule:

$$\begin{aligned} \tau'' &= \tau' + \tau + i\,\tilde{\chi}'_s\,\tilde{\chi}_s \,, \\ \mu''_s &= \mu'_s + \mu_s + i\,\tilde{\chi}'_s\,\tilde{\chi}_s \,, \qquad \text{(no sum over } s)\,, \\ \tilde{\chi}''_s &= \tilde{\chi}'_s + \tilde{\chi}_s \,. \end{aligned} \tag{11.77}$$

which gives the group transformation rule given in Eq. (11.74). We now can extend superspace to include an additional real parameter x_r, which transform according to the rule:

$$\begin{aligned} t' &= t + \tau + i\tilde{\chi}_s\,\tilde{\theta}_s \,, \\ x'_s &= x_s + \mu_s + i\,\tilde{\chi}_s\,\tilde{\theta}_s \,, \qquad \text{(no sum over } s)\,, \\ \tilde{\theta}'_s &= \tilde{\theta}_s + \tilde{\chi}_s \,. \end{aligned} \tag{11.78}$$

Superspace is now described by $1 + 2N$ coordinates: $\tilde{s} = (\,t, \tilde{\theta}_s, x_s\,)$. This redefinition of the coordinates reproduces the supersymmetry transformation with the central charges included as group operators. The $1 + 2N$ generators of the group are now defined by the infinitesimal unitary transformation:

$$U_{\rm ext}(1 + \Delta\tilde{g}) = 1 + i\,\Delta\tau\, H + \Delta\tilde{\chi}_s\,\tilde{Q}_s + \Delta\mu_s\, M_s + \cdots \,, \tag{11.79}$$

which now incorporates the projective phase factor in the extended unitary transformation. The group transformation rule now reads:

$$U_{\rm ext}(\tilde{g}')\, U_{\rm ext}(\tilde{g}) = U_{\rm ext}(\tilde{g}'\tilde{g}) \,, \tag{11.80}$$

with *no* phase factor. The extended group algebra is the same as that given in Theorem 32 with $U(g)$ replaced by $U_{\rm ext}(\tilde{g})$.

In order to define transformations of Hilbert space operators, other than the generators of the supersymmetry transformation, it is *necessary* to extend the group; otherwise we have no idea how to incorporate the important phase factor (central charges) for transformations of superfunctions of the superspace variables.

11.7 Differential forms

In this section, we find the transformations of differential forms under the extended $\mathcal{SUSY}_{\rm ext}$ transformations, the differential forms $(\,\mathrm{d}t, \mathrm{d}\tilde{\theta}_s, \mathrm{d}x_s\,)$ transform according to:

$$\begin{aligned} \mathrm{d}t' &= \mathrm{d}t + \tau + i\,\tilde{\chi}_s\,\mathrm{d}\tilde{\theta}_s\,, \\ \mathrm{d}x'_s &= \mathrm{d}x_s + \mu_s + i\,\tilde{\chi}_s\,\mathrm{d}\tilde{\theta}_s\,, \qquad (\text{no sum over } s)\,, \\ \mathrm{d}\tilde{\theta}'_s &= \mathrm{d}\tilde{\theta}_s + \tilde{\chi}_s\,. \end{aligned} \tag{11.81}$$

Derivatives transform in the opposite way. The inverse of Eq. (11.78) is given by:

$$\begin{aligned} t &= t' - \tau - i\tilde{\chi}_s\,\tilde{\theta}_s\,, \\ x_s &= x'_s - \mu_s - i\,\tilde{\chi}_s\,\tilde{\theta}_s\,, \qquad (\text{no sum over } s)\,, \\ \tilde{\theta}_s &= \tilde{\theta}'_s - \tilde{\chi}_s\,. \end{aligned} \tag{11.82}$$

So we find:

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial t}\,, \qquad \frac{\partial}{\partial x'_s} = \frac{\partial}{\partial x_s}\,, \qquad \frac{\partial}{\partial \tilde{\theta}'_s} = \frac{\partial}{\partial \tilde{\theta}'_s} + i\,\tilde{\chi}_s\left[\,\frac{\partial}{\partial t} + \frac{\partial}{\partial x_s}\,\right]. \tag{11.83}$$

In the last equation, there is no sum over s. In a short-hand notation, we write these equations as:

$$\partial'_t = \partial_t\,, \qquad \partial'_{x_s} = \partial_{x_s}\,, \qquad \partial'_{\tilde{\theta}_s} = \partial_{\tilde{\theta}_s} + i\,\tilde{\chi}_s\,[\,\partial_t + \partial_{x_s}\,]\,. \tag{11.84}$$

An invariant superderivative is now defined by:

$$D_s = \partial_{\tilde{\theta}_s} - i\,\tilde{\theta}_s\,[\,\partial_t + \partial_{x_s}\,]\,, \tag{11.85}$$

where, again, there is no sum over s in the last term. D_s is constructed to be invariant:

$$\begin{aligned} D'_s &= \partial'_{\tilde{\theta}_s} - i\,\tilde{\theta}'_s\,[\,\partial'_t + \partial'_{x_s}\,] \\ &= \partial_{\tilde{\theta}_s} - i\,\tilde{\chi}_s\,[\,\partial_t + \partial_{x_s}\,] + i\,[\,\tilde{\theta}_s + \tilde{\chi}_s\,]\,[\,\partial_t + \partial_{x_s}\,] \\ &= \partial_{\tilde{\theta}_s} - i\,\tilde{\theta}_s\,[\,\partial_t + \partial_{x_s}\,] = D_s\,, \end{aligned} \tag{11.86}$$

as desired.

We also find:

$$\{\,D_s, D_{s'}\,\} = -2i\,\delta_{ss'}\,[\,\partial_t + \partial_{x_s}\,]\,. \tag{11.87}$$

References

[1] V. Bargmann, "On unitary ray representations of continuous groups," Ann. Math. **59**, 1 (1954).

Part II

Applications

Chapter 12

Finite quantum systems

In this chapter, we discuss quantum systems that can be described by finite matrices. As an example of such systems, we study electrons that are confined to be located on a finite number of fixed atomic sites, such as molecules. We study diatomic molecules, periodic and liner chains. Electrons on a lattice are discussed further in Chapter 17.

12.1 Diatomic molecules

In this section, we discuss some toy molecular systems consisting of a finite number of "atoms," with electrons free to move between the atomic cites.

The general approach for studying molecules is one that was developed by Born and Oppenheimer, and goes by that name. Since the atoms are much heavier than the electrons, the electron motion is first solved assuming that the atoms are at rest. The attractive potential created by the electrons and the repulsive potential between the atomic centers, then provide an overall potential which can bind the atoms. The atoms then execute small vibrations about the equilibrium seperation distance.

We start by considering diatomic molecules, consisting of two identical atomic sites, labeled $|\,1\,\rangle$ and $|\,2\,\rangle$, and an electron which can jump from one site to the other. The state of the electron at any time t is then written as:

$$|\,q(t)\,\rangle = q_1(t)\,|\,1\,\rangle + q_2(t)\,|\,2\,\rangle\,, \tag{12.1}$$

where $q_1(t)$ and $q_2(t)$ are the amplitudes of finding the electron at sites 1 and 2 respectively at time t. The wave function of the electron is given approximately by:[1]

$$\psi(x,t) = q_1(t)\,\psi_1(x) + q_2(t)\,\psi_2(x)\,, \qquad \psi_1(x) = \psi_0(x-a/2)\,, \qquad \psi_2(x) = \psi_0(x+a/2)\,, \tag{12.2}$$

where $\psi_{1,2}(x)$ are the ground state wave functions for the electron for the isolated atoms, which we take to be the same, with energy ϵ_0. The potential energy as seen by an electron is illustrated in Fig. 12.1. Considered as a two-state system, the Hamiltonian for the electron is given in matrix form by:

$$H = \begin{pmatrix} \epsilon_0 & -\Gamma_0 \\ -\Gamma_0 & \epsilon_0 \end{pmatrix}\,, \tag{12.3}$$

where $\Gamma_0/\hbar$ is the transition rate for the electron to move between sites. Referring to Fig. 12.1, Γ_0 is given by the overlap integral:

$$\Gamma_0 = -\int_{-\infty}^{+\infty} \psi_1^*(x)\,\big[\,V_1(x) + V_2(x)\,\big]\,\psi_2(x)\,\mathrm{d}x > 0\,, \tag{12.4}$$
$$V_1(x) = V_0(x-a/2)\,, \qquad V_2(x) = V_0(x+a/2)\,,$$

[1] We discuss the diatomic molecule in much greater detail and more accuracy in Appendix ZZ.

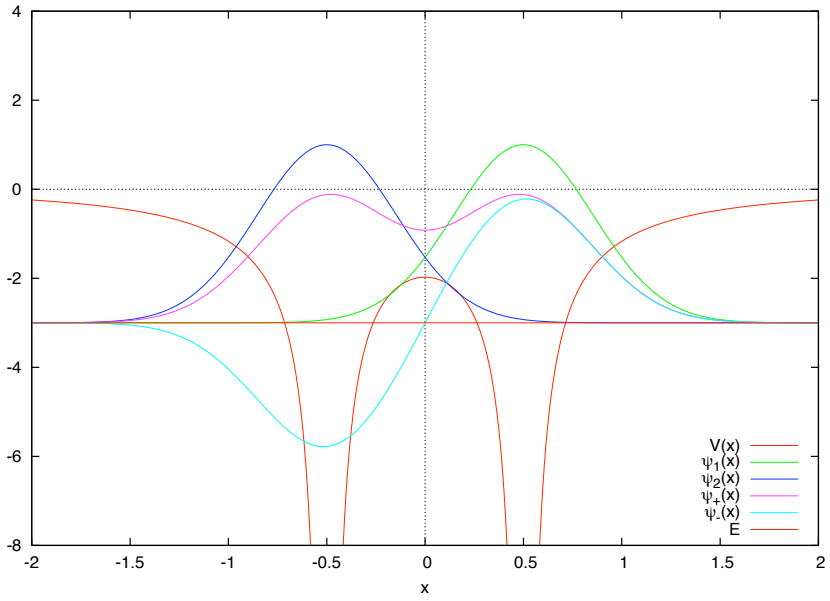

Figure 12.1: We plot the potential energy for an electron in two atomic sites. We also sketch wave functions $\psi_{1,2}(x)$ for an electron in the isolated atomic sites and the symmetric and antisymmetric combinations $\psi_{\pm}(x)$.

since the potential energy in the overlap region is *negative*. The wave functions are real, so Γ_0 is also real. The eigenvalue problem for the energy,

$$H\,|\,E_n\,\rangle = E_n\,|\,E_n\,\rangle\,, \tag{12.5}$$

has the solutions:

$$\begin{aligned} |\,E_+\,\rangle &= \frac{1}{\sqrt{2}}\,(\,|\,1\,\rangle + |\,2\,\rangle\,)\,, \qquad E_+ = \epsilon_0 - \Gamma_0\,, \\ |\,E_-\,\rangle &= \frac{1}{\sqrt{2}}\,(\,|\,1\,\rangle - |\,2\,\rangle\,)\,, \qquad E_- = \epsilon_0 + \Gamma_0\,. \end{aligned}$$

The state of an electron at any time t is given by:

$$\begin{aligned} |\,q(t)\,\rangle &= q_+\,e^{-iE_+t/\hbar}\,|\,E_+\,\rangle + q_-\,e^{-iE_-t/\hbar}\,|\,E_-\,\rangle \\ &= q_1(t)\,|\,1\,\rangle + q_2(t)\,|\,2\,\rangle\,, \end{aligned} \tag{12.6}$$

where

$$\begin{aligned} q_1(t) &= \frac{e^{-i\epsilon_0 t/\hbar}}{\sqrt{2}}\,\left(\,q_+\,e^{+i\Gamma_0 t/\hbar} + q_-\,e^{-i\Gamma_0 t/\hbar}\,\right) = e^{-i\epsilon_0 t/\hbar}\,\left(\,q_1\,\cos(\Gamma_0 t/\hbar) + q_2\,\sin(\Gamma_0 t/\hbar)\,\right)\,, \\ q_2(t) &= \frac{e^{-i\epsilon_0 t/\hbar}}{\sqrt{2}}\,\left(\,q_+\,e^{+i\Gamma_0 t/\hbar} - q_-\,e^{-i\Gamma_0 t/\hbar}\,\right) = e^{-i\epsilon_0 t/\hbar}\,\left(\,q_1\,\sin(\Gamma_0 t/\hbar) + q_2\,\cos(\Gamma_0 t/\hbar)\,\right)\,, \end{aligned} \tag{12.7}$$

and where $q_\pm = (\,q_1 \mp i\,q_2\,)/\sqrt{2}$ are fixed by the initial conditions.

So the ground state of the electron $|\,E_+\,\rangle$ is the *even* parity or symmetric combination. In this state the electron is found with higher probability between the two atomic centers. The excited state $|\,E_-\,\rangle$ is an *odd* parity or antisymmetric state, with the electron found with higher probability outside of the two atomic

centers. We sketch the wave functions $\psi_{\pm}(x)$ for the eigenstates in Fig. 12.1. The transition rate $\Gamma_0/\hbar$ increases with decreasing separation of the atoms, whereas the repulsive force between the atomic centers increases with decreasing separation. The balancing of these two forces provides a potential which can bind the molecule. Since two electrons can be put in one orbital state with paired spins, molecules with paired electrons should have stronger binding. In Appendix ZZ, we give a variational calculation of the potential energy between the atoms in the Hydrogen molecule. A plot of the potential function is given in Fig. XX. The binding energy and bond length are given by $E = 0.0$ and $a = 0.744\,\text{Å}$ for H_2.

Exercise 22. Suppose the electron in a diatomic molecule can be in the first excited state ϵ_1 with *odd* parity wave functions $\psi_1(x \pm a/2)$ of the atoms at sites $|\,1\,\rangle$ or $|\,2\,\rangle$.

1. Sketch the wave functions for these two sites and find an integral for the transition rate $\Gamma_1/\hbar$ for the electron to jump between the two sites so that it remains in the excited states. Is Γ_1 positive or negative?
2. Write down the Hamiltonian for this problem, assuming no mixing transitions between the ground and excited states. Find the eigenvalues and eigenvectors of the Hamiltonian operator, and list the eigenvalues in increasing order.
3. Suppose the electron can jump between the two levels with a rate $\Gamma_{01}/\hbar$ and $\Gamma_{10}/\hbar$. What are the signs and relative magnetudes of $\Gamma_{01}/\hbar$ and $\Gamma_{10}/\hbar$? Write down the Hamiltonian for this case, but do not solve.

Exercise 23. Consider the "molecule" consisting of four sites, as shown in Fig. 12.2 below. The energy of the electron at each site is given by ϵ_0 and the transition rates between sites connected by a solid line are all equal to $\Gamma_0/\hbar$. Using a basis set $|\,1\,\rangle, |\,2\,\rangle, |\,3\,\rangle, |\,4\,\rangle$ for each site, the Hamiltonian is given by the matrix:

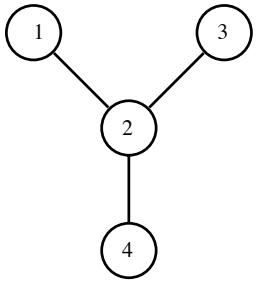

Figure 12.2: A molecule containing four atoms.

$$H = \begin{pmatrix} \epsilon_0 & -\Gamma_0 & 0 & 0 \\ -\Gamma_0 & \epsilon_0 & -\Gamma_0 & -\Gamma_0 \\ 0 & -\Gamma_0 & \epsilon_0 & 0 \\ 0 & -\Gamma_0 & 0 & \epsilon_0 \end{pmatrix} \tag{12.8}$$

1. Show that the eigenvalues and eigenvectors are given by:

$$|\,E_1\,\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ 0 \\ -1 \\ -1 \end{pmatrix}, \qquad E_1 = \epsilon_0\,, \qquad |\,E_3\,\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ \sqrt{3} \\ 1 \\ 1 \end{pmatrix}, \qquad E_3 = \epsilon_0 - \sqrt{3}\,\Gamma_0\,,$$

$$|\,E_2\,\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \qquad E_2 = \epsilon_0\,, \qquad |\,E_4\,\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -\sqrt{3} \\ 1 \\ 1 \end{pmatrix}, \qquad E_1 = \epsilon_0 + \sqrt{3}\,\Gamma_0\,.$$

2. If at $t = 0$ the electron is located at site $|\,2\,\rangle$, find the ket $|\,q(t)\,\rangle$ for all time and show that the probability of finding the electron on site $|\,1\,\rangle$ as a function of time is given by:

$$P_1(t) = |\,\langle\,1\,|\,q(t)\,\rangle\,|^2 = \frac{1}{3}\,\sin^2(\,\sqrt{3}\Gamma_0 t/\hbar\,)\,. \tag{12.9}$$

12.2 Periodic chains

The dynamics of an electron jumping between circular or periodic chains of N atomic sites, such as the benzine molecule (C_6H_6), can be treated in much the same way as in the last section. Consider the arrangement, for example, of $N = 6$ atoms as shown in Fig. 12.3. We describe the electron again by the ket:

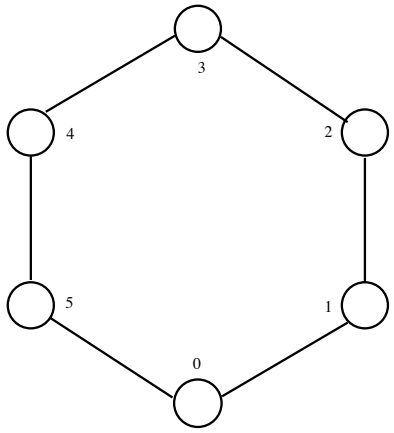

Figure 12.3: A molecule containing six atomic sites, arranged in a circular chain.

$$|\,q(t)\,\rangle = \sum_{n=0}^{N-1} q_n(t)\,|\,n\,\rangle\,. \tag{12.10}$$

The periodic requirement means that $|\,0\,\rangle = |\,N\,\rangle$. Because of this periodic requirement, it will be useful to change basis sets to a new basis by using a finite Fourier transform. Let us define this new basis $\widetilde{|\,k\,\rangle}$ by the set of equations:

$$\begin{aligned} |\,n\,\rangle &= \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} e^{+2\pi i\,kn/N}\,\widetilde{|\,k\,\rangle}\,, \\ \widetilde{|\,k\,\rangle} &= \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} e^{-2\pi i\,kn/N}\,|\,n\,\rangle\,. \end{aligned} \tag{12.11}$$

Then $|\,0\,\rangle = |\,N\,\rangle$, as required, and

$$\langle\,n\,|\,n'\,\rangle = \delta_{n,n'}\,, \qquad \widetilde{\langle\,k\,|}\;\widetilde{|\,k'\,\rangle} = \delta_{k,k'}\,. \tag{12.12}$$

The Hamiltonian in the $|\,n\,\rangle$ basis is given by:

$$H = \sum_{n=0}^{N-1}\Big\{\,\epsilon_0\,|\,n\,\rangle\langle\,n\,| - \Gamma_0\,\big[\,|\,n\,\rangle\langle\,n+1\,| + |\,n+1\,\rangle\langle\,n\,|\,\big]\,\Big\}\,. \tag{12.13}$$

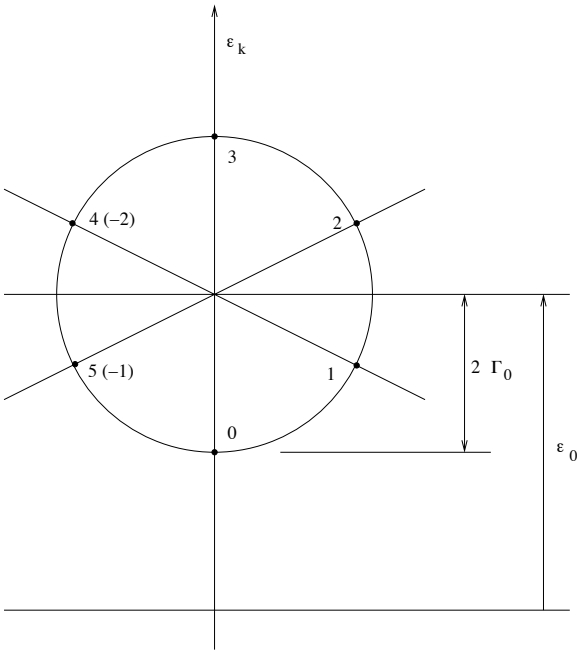

Figure 12.4: Construction for finding the six eigenvalues for an electron on the six periodic sites of Fig. 12.3, for values of $k = 0, \ldots, 5$. Note the degeneracies for values of $k = 1, 5$ and $k = 2, 4$.

Now we have:

$$\begin{aligned}
\sum_{n=0}^{N-1} |\, n \,\rangle\langle\, n \,| &= \sum_{k=0}^{N-1} \widetilde{|\, k \,\rangle\langle\, k \,|}\,, \\
\sum_{n=0}^{N-1} \left[|\, n \,\rangle\langle\, n+1 \,| + |\, n+1 \,\rangle\langle\, n \,| \right] &= \sum_{k=0}^{N-1} \left[e^{+2\pi i\, k/N} + e^{-2\pi i\, k/N} \right] \widetilde{|\, k \,\rangle\langle\, k \,|}\,, \\
&= \sum_{k=0}^{N-1} 2\cos(\, 2\pi k/N) \, \widetilde{|\, k \,\rangle\langle\, k \,|}\,,
\end{aligned} \tag{12.14}$$

So in the Fourier transform basis, the Hamiltonian becomes

$$H = \sum_{k=0}^{N-1} \epsilon_k \, \widetilde{|\, k \,\rangle\langle\, k \,|}\,, \qquad \epsilon_k = \epsilon_0 - 2\Gamma_0 \cos(\, 2\pi k/N)\,, \tag{12.15}$$

and is diagonal. Solutions to the eigenvalue problem,

$$H \,|\, E_k \,\rangle = E_k \,|\, E_k \,\rangle\,,$$

are easy in this basis. We find $E_k = \epsilon_k$ and $|\, E_k \,\rangle = \widetilde{|\, k \,\rangle}$, for $k = 0, \ldots, N-1$. A construction for finding the eigenvalues is shown for the case when $N = 6$ in Fig. 12.4. Note the degeneracies for $k = 1, 5$ and $k = 2, 6$. For this reason, it is useful to map $k = 5$ to $k = -1$ and $k = 4$ to $k = -2$. Then the range of k is $-2 \le k \le +3$, with the states degenerate for $\pm k$ and the eigenvalues functions of $|k|$ only.

12.3 Linear chains

For linear chains, we can apply methods similar to that used for periodic sites, with different boundary conditions. Consider, for example, the case of $N = 6$ atoms arranged in a linear chain, as shown in Fig. 12.5. Here we label each site by the kets: $|\, n\,\rangle$, $n = 1, \ldots, N$ so that the state of an electron at time t is given by:

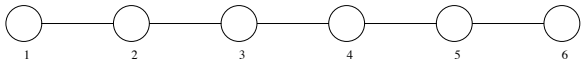

Figure 12.5: A molecule containing six atomic sites, arranged in a linear chain.

$$|\, q(t)\,\rangle = \sum_{n=1}^{N} q_n(t)\,|\, n\,\rangle\,. \tag{12.16}$$

If we extend the sites to include $n = 0$ and $n = N + 1$ and require that $|\, 0\,\rangle = |\, N + 1\,\rangle = 0$, then we can expand in a finite Fourier sine transform:

$$\begin{aligned} |\, n\,\rangle &= \frac{1}{\sqrt{N+1}} \sum_{k=1}^{N} \sin(\pi nk/(N+1))\,\widetilde{|\, k\,\rangle}\,, \\ \widetilde{|\, k\,\rangle} &= \frac{1}{\sqrt{N+1}} \sum_{n=1}^{N} \sin(\pi nk/(N+1))\,|\, n\,\rangle\,, \end{aligned} \tag{12.17}$$

which satisfies the required boundary conditions. Note that $\widetilde{|\, 0\,\rangle} = \widetilde{|\, N+1\,\rangle} = 0$ also. The Hamiltonian is again given by:

$$H = \sum_{n=1}^{N} \Big\{ \epsilon_0\,|\, n\,\rangle\langle\, n\,| - \Gamma_0 \left[\, |\, n\,\rangle\langle\, n+1\,| + |\, n+1\,\rangle\langle\, n\,|\,\right] \Big\}\,. \tag{12.18}$$

which differs from (12.13) only by the range of n. Again, we find:

$$\begin{aligned} \sum_{n=1}^{N} |\, n\,\rangle\langle\, n\,| &= \sum_{k=1}^{N} \widetilde{|\, k\,\rangle}\widetilde{\langle\, k\,|}\,, \\ \sum_{n=1}^{N} \left[\, |\, n\,\rangle\langle\, n+1\,| + |\, n+1\,\rangle\langle\, n\,|\,\right] &= \sum_{k=1}^{N} 2\cos(\pi k/(N+1))\,\widetilde{|\, k\,\rangle}\widetilde{\langle\, k\,|}\,, \end{aligned} \tag{12.19}$$

So in the Fourier sine transform basis, the Hamiltonian becomes

$$H = \sum_{k=1}^{N} \epsilon_k\,\widetilde{|\, k\,\rangle}\widetilde{\langle\, k\,|}\,, \qquad \epsilon_k = \epsilon_0 - 2\Gamma_0\cos(\pi k/(N+1))\,, \tag{12.20}$$

and is diagonal, as we found in the periodic case. Solutions to the eigenvalue problem,

$$H\,|\, E_k\,\rangle = E_k\,|\, E_k\,\rangle\,,$$

are again simple in this basis. We find $E_k = \epsilon_k$ and $|\, E_k\,\rangle = \widetilde{|\, k\,\rangle}$, for $k = 1, \ldots, N$. A construction for finding the eigenvalues for a linear chain is shown for the case when $N = 6$ in Fig. 12.6. There are no degeneracies in this case. Eigenvectors for the linear chain are standing waves for the electron on the lattice rather than the travelling waves for the periodic lattice. Since standing waves can be constructed from the superposition of two travelling waves, which on the periodic lattice are degenerate, the eigenvalues for an electron on a linear chain of atoms are similar to the eigenvalues for a periodic chain of atoms. Because of the boundary conditions, the $k = 0$ and $k = N + 1$ mode is missing, and the angular spacing of the eigenvalues in the construction is half that of a periodic chain with $N + 1$ atoms.

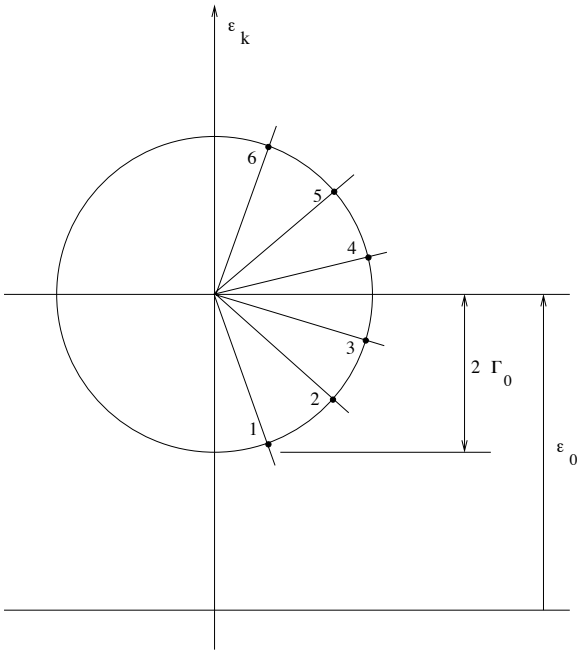

Figure 12.6: Six eigenvalues for the six linear sites of Fig. 12.5, for values of $k = 1, \ldots, 6$.

12.4 Impurities

Consider a long chain of N atoms with an impurity atom located at site $n = 0$, as shown in Fig. 12.7. Again

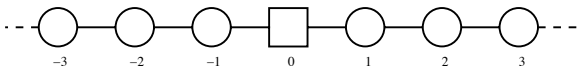

Figure 12.7: A long chain with an impurity atom at site 0.

assuming nearest neighbor interactions only, we can write the Hamiltonian for this system as:

$$H = \sum_{n=1}^{N-1} \Big\{ \epsilon_0 \left[\,|\, n \,\rangle\langle\, n \,|\, + |\, -n \,\rangle\langle\, -n \,|\, \right] - \Gamma_0 \left[\,|\, n \,\rangle\langle\, n+1 \,|\, + |\, n+1 \,\rangle\langle\, n \,|\, + |\, -n-1 \,\rangle\langle\, -n \,|\, + |\, -n \,\rangle\langle\, -n-1 \,|\, \right] \Big\}$$

s

$$+ \epsilon_1 \,|\, 0 \,\rangle\langle\, 0 \,|\, - \Gamma_1 \left[\,|\, 0 \,\rangle\langle\, 1 \,|\, + |\, 1 \,\rangle\langle\, 0 \,|\, + |\, -1 \,\rangle\langle\, 0 \,|\, + |\, 0 \,\rangle\langle\, -1 \,|\, \right] . \quad (12.21)$$

12.4.1 Bound state

Under certain conditions, the electron can become trapped at the $n = 0$ site. We study the bound states in this section. The eigenvalue equation for bound states is:

$$H \,|\, \psi_E \,\rangle = E \,|\, \psi_E \,\rangle \,. \quad (12.22)$$

We expanding the eigenvector in the form:

$$|\,\psi_E\,\rangle = \sum_{n=-\infty}^{+\infty} q_n\,|\,n\,\rangle\,. \tag{12.23}$$

Substitution into Eq. (12.22) and using (12.21) gives the three equations for the coefficients q_n for $n = 0, \pm 1$:

$$\begin{aligned} -\Gamma_1\,q_1 + \epsilon_1\,q_0 - \Gamma_1\,q_{-1} &= E\,q_0\,,\\ -\Gamma_0\,q_2 + \epsilon_0\,q_1 - \Gamma_1\,q_0 &= E\,q_1\,,\\ -\Gamma_1\,q_0 + \epsilon_0\,q_{-1} - \Gamma_0\,q_{-2} &= E\,q_{-1}\,, \end{aligned} \tag{12.24}$$

and for $n > 1$ and $n < -1$, we find:

$$-\,\Gamma_0\,q_{n+1} + \epsilon_0\,q_n - \Gamma_0\,q_{n-1} = E\,q_n\,. \tag{12.25}$$

Assuming a solution of the form:

$$q_n = \begin{cases} q_+\,e^{-n\theta}\,, & \text{for } n > 0,\\ q_-\,e^{+n\theta}\,, & \text{for } n < 0. \end{cases} \tag{12.26}$$

Our task is to find q_0, $q_\pm$, θ, and E. Eq. (12.25) is satisfied for $n > 1$ and $n < -1$ if:

$$\epsilon_0 - E = 2\Gamma_0\,\cosh(\theta) = \Gamma_0\left[\,e^{+\theta} + e^{-\theta}\,\right]. \tag{12.27}$$

Eqs. (12.24) are satisfied if:

$$\begin{aligned} (\,\epsilon_1 - E\,)\,q_0 &= \Gamma_1\,(\,q_+ + q_-\,)\,e^{-\theta}\,,\\ (\,\epsilon_0 - E - \Gamma_0\,e^{-\theta}\,)\,q_+ &= \Gamma_1\,q_0\,e^{\theta}\,,\\ (\,\epsilon_0 - E - \Gamma_0\,e^{-\theta}\,)\,q_- &= \Gamma_1\,q_0\,e^{\theta}\,. \end{aligned} \tag{12.28}$$

Solving for q_+ and q_-, we find:

$$q_+ = q_- = \frac{\Gamma_1\,e^{\theta}}{\epsilon_0 - E - \Gamma_0\,e^{-\theta}}\,q_0 = \frac{\Gamma_1}{\Gamma_0}\,q_0\,. \tag{12.29}$$

so that $q_+ = q_-$. Then the first of Eqs. (12.28) gives:

$$\epsilon_1 - E = 2\,\frac{\Gamma_1^2}{\Gamma_0}\,e^{-\theta}\,. \tag{12.30}$$

Combining Eqs. (12.27) and (12.30) gives a transcendental equation for θ:

$$(\epsilon_0 - \epsilon_1)/\Gamma_0 = e^{\theta} + \left[\,1 - 2\left(\Gamma_1/\Gamma_0\right)^2\,\right]e^{-\theta}\,. \tag{12.31}$$

In Fig. XX, we show a plot of the right and left sides of this equation for the case when $(\epsilon_0 - \epsilon_1)/\Gamma_0 = 0.2667$ and $\Gamma_1/\Gamma_0 = 0.8$. For this case, we find $\theta = YY$, from which we can find the energy eigenvalue E. There is only one bound state.

The eigenvector for the bound state is given by:

$$|\,\psi_E\,\rangle = q_0\left\{\sum_{n=1}^{\infty}\frac{\Gamma_1}{\Gamma_0}\,\left[\,e^{-\theta\,n}\,|\,n\,\rangle + e^{+\theta\,n}\,|\,-n\,\rangle\,\right] + |\,0\,\rangle\right\}, \tag{12.32}$$

where q_0 is a normalization factor.

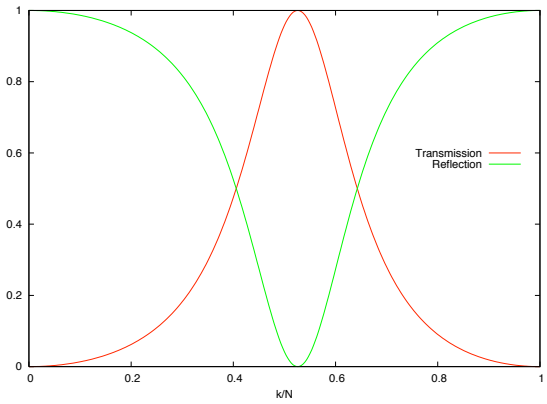

Figure 12.8: Transmission and reflection coefficients for electron scattering from an impurity for the case when $(\epsilon_0 - \epsilon_1)/\Gamma_0 = 0.2667$ and $\Gamma_1/\Gamma_0 = 0.8$.

12.4.2 Scattering

An electron can also scatter off the impurity site at $n = 0$. In order to study the transmission and reflection from the impurity, let us find electron solutions for a fixed value of energy E_k to the left and right of the impurity. This vector is given by:

$$|\,\psi_k\,\rangle = \sum_{n=1}^{N-1}\Big\{\,\big[\,A_k\,e^{-i\theta_k\,n} + B_k\,e^{+i\theta_k\,n}\,\big]\,|\,-n\,\rangle + \big[\,C_k\,e^{+i\theta_k\,n} + D_k\,e^{-i\theta_k\,n}\big]\,|\,n\,\rangle\,\Big\} + q_0\,|\,0\,\rangle\,. \tag{12.33}$$

Here $H\,|\,\psi_k\,\rangle = E_k\,|\,\psi_k\,\rangle$, where

$$E_k = \epsilon_0 - 2\Gamma_0\,\cos(\theta_k)\,, \qquad \text{with} \qquad \theta_k = \frac{2\pi k}{N}\,, \tag{12.34}$$

and the Hamiltonian H is given by Eq. (12.21). Our task is to find the relation of the "out" states B_k, C_k to the "in" states, A_k, D_k, and to find q_0. Computing the overlaps:

$$\langle\,n\,|\,H\,|\,\psi_k\,\rangle = E_k\,\langle\,n\,|\,\psi_k\,\rangle\,, \tag{12.35}$$

we see that this equation is identically satisfied for $n > 1$ and $n < -1$. For $n = 0, \pm 1$, we find the equations:

$$(\,\epsilon_1 - E_k\,)\,q_0 = \Gamma_1\,\big[\,A_k e^{-i\theta_k} + B_k e^{+i\theta_k} + C_k e^{+i\theta_k} + D_k e^{-i\theta_k}\,\big] \tag{12.36}$$

$$\big[\,\epsilon_0 - E_k - \Gamma_0\,e^{-i\theta_k}\,\big]\,A_k\,e^{-i\theta_k} + \big[\,\epsilon_0 - E_k - \Gamma_0\,e^{+i\theta_k}\,\big]\,B_k\,e^{+i\theta_k} = \Gamma_1\,q_0 \tag{12.37}$$

$$\big[\,\epsilon_0 - E_k - \Gamma_0\,e^{+i\theta_k}\,\big]\,C_k\,e^{+i\theta_k} + \big[\,\epsilon_0 - E_k - \Gamma_0\,e^{-i\theta_k}\,\big]\,D_k\,e^{-i\theta_k} = \Gamma_1\,q_0 \tag{12.38}$$

Now from (12.34), we find:

$$\epsilon_0 - E_k - \Gamma_0\,e^{\pm i\theta_k} = \Gamma_0\,e^{\mp i\theta_k}\,, \tag{12.39}$$

so (12.37) and (12.38) become:

$$\begin{aligned} A_k + B_k &= (\,\Gamma_1/\Gamma_0\,)\,q_0\,, \\ C_k + D_k &= (\,\Gamma_1/\Gamma_0\,)\,q_0\,, \end{aligned} \tag{12.40}$$

whereas (12.36) becomes:

$$\big[\,(\,\epsilon_1 - \epsilon_0\,)/\Gamma_0 - 2\cos(\theta_k)\,\big]\,q_0 = (\,\Gamma_1/\Gamma_0\,)\,\big[\,(\,A_k + D_k\,)\,e^{-i\theta_k} + (\,B_k + C_k\,)\,e^{+i\theta_k}\,\big] \tag{12.41}$$

So from (12.40) and (12.41), we find the equations:

$$\begin{aligned} A_k + B_k &= -\beta_k \left[\left(A_k + D_k \right) e^{-i\theta_k} + \left(B_k + C_k \right) e^{+i\theta_k} \right] , \\ C_k + D_k &= -\beta_k \left[\left(A_k + D_k \right) e^{-i\theta_k} + \left(B_k + C_k \right) e^{+i\theta_k} \right] , \end{aligned} \tag{12.42}$$

where

$$\beta_k = \left[\frac{\Gamma_1}{\Gamma_0} \right]^2 \frac{1}{\left(\epsilon_0 - \epsilon_1 \right)/\Gamma_0 + 2\cos(\theta_k)} \, . \tag{12.43}$$

Eqs. (12.42) can be written as:

$$\begin{aligned} \left[1 + \beta_k \, e^{-i\theta_k} \right] A_k + \left[1 + \beta_k \, e^{+i\theta_k} \right] B_k + \beta_k \, e^{+i\theta_k} \, C_k + \beta_k \, e^{-i\theta_k} \, D_k = 0 \, , \\ \beta_k \, e^{-i\theta_k} \, A_k + \beta_k \, e^{+i\theta_k} \, B_k + \left[1 + \beta_k \, e^{+i\theta_k} \right] C_k + \left[1 + \beta_k \, e^{-i\theta_k} \right] D_k = 0 \, , \end{aligned} \tag{12.44}$$

from which we can find solutions of the out states (B_k, C_k) in terms of the in states (A_k, D_k):

$$\begin{pmatrix} B_k \\ C_k \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A_k \\ D_k \end{pmatrix} \, , \tag{12.45}$$

with

$$S_{11} = S_{22} = -\frac{1 + 2\beta_k \cos(\theta_k)}{1 + 2\beta_k \, e^{+i\theta_k}} \, , \qquad S_{12} = S_{21} = -\frac{2i\beta_k \sin(\theta_k)}{1 + 2\beta_k \, e^{+i\theta_k}} \, . \tag{12.46}$$

Unitarity of the S matrix requires that

$$S^\dagger S = \begin{pmatrix} S_{11}^* & S_{21}^* \\ S_{12}^* & S_{22}^* \end{pmatrix} \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} |S_{11}|^2 + |S_{12}|^2 & S_{12}^* S_{11} + S_{22}^* S_{21} \\ S_{11}^* S_{12} + S_{21}^* S_{22} & |S_{21}|^2 + |S_{22}|^2 \end{pmatrix} = 1 \, . \tag{12.47}$$

It is easy to check that the solutions (12.46) satisfy the unitary relations (12.47). Transmission (T) and reflection (R) coefficients are given by:

$$T = |S_{12}|^2 = |S_{21}|^2 \, , \qquad R = |S_{11}|^2 = |S_{22}|^2 \, , \tag{12.48}$$

and are plotted in Fig. 12.8 for the case when $(\epsilon_0 - \epsilon_1)/\Gamma_0 = 0.2667$ and $\Gamma_1/\Gamma_0 = 0.8$, as a function of k/N.

Exercise 24. Find the transmission and reflection coefficents for scattering of an electron from two long lines of atoms connected at site $n = 0$, as shown in Fig 12.9. The electron energies and jumping rates for atoms for $n < 0$ are given by ϵ_0 and $\Gamma_0/\hbar$, whereas for $n \geq 0$, the energies and jumping rates are given by ϵ_1 and $\Gamma_1/\hbar$. Take the jumping rate between the $n = -1$ and $n = 0$ sites as Γ_{10}.

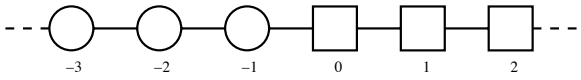

Figure 12.9: Two long connected chains.

References

Chapter 13

One and two dimensional wave mechanics

13.1 Introduction

Quantum wires are systems with one continuous coordinate dimension x. Conducting thin-films are examples of quantum systems in two space dimensions. These low-dimension systems can frequently be solved exactly, using Schrödinger's equation, and are useful for understanding systems in higher dimension, where computation can be more difficult. But one and two dimensional systems are interesting in their own right. We discuss first quantum mechanics in one dimension.

13.2 Schrödinger's equation in one dimension

Schrödinger's equation in one dimension is:

$$\left\{ -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \right\} \psi(x,t) = i\hbar \frac{\partial \psi(x,t)}{\partial t} \,. \tag{13.1}$$

Probability conservation: This equation obeys a conservation equation:

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial j(x,t)}{\partial x} = 0 \,, \tag{13.2}$$

where

$$\rho(x,t) = |\psi(x,t)|^2 \,, \qquad j(x,t) = \frac{\hbar}{2mi}\left\{ \psi^*(x,t)\frac{\partial \psi(x,t)}{\partial x} - \frac{\partial \psi^*(x,t)}{\partial x}\psi(x,t) \right\} \,. \tag{13.3}$$

Time reversal: Schrödinger's equation is also invariant under time-reversal. Reversing the time variable in Eq. (13.1) gives:

$$\left\{ -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \right\} \psi(x,-t) = -i\hbar \frac{\partial \psi(x,-t)}{\partial t} \,. \tag{13.4}$$

Now take the complex conjugate. Since $V(x)$ is real, we have:

$$\left\{ -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \right\} \psi^*(x,-t) = i\hbar \frac{\partial \psi^*(x,-t)}{\partial t} \,, \tag{13.5}$$

which is the same equation we started with. So if $\psi(x,t)$ is a solution of Schrödinger's equation, then $\psi^*(x,-t)$ is a solution also. If we separate variables according to:

$$\psi(x,t) = \int \frac{\mathrm{d}k}{2\pi}\,\psi_k(x)\,e^{-iE_kt/\hbar}\,, \qquad E_k = \frac{\hbar^2\,k^2}{2m}\,, \tag{13.6}$$

then time-reversal invariance means that if $\psi_k(x)$ is a solution of the time-independent Schrödinger equation, then $\psi_k^*(x)$ is also.

Parity: Parity is reversal of the x coordinate. If $V(-x) = V(x)$, then if $\psi(x,t)$ is a solution of Schrödinger's equation, then $\psi(-x,t)$ is also a solution. We will use these conservation and symmetry relations in this chapter.

13.2.1 Transmission of a barrier

In this section, we discuss transmission of particles by a barrier of general shape, as shown in Fig. XX. Schrödinger's time-independent equation for this problem is given by:

$$\left\{-\frac{\hbar^2}{2m}\,\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right\}\,\psi_k(x) = \frac{\hbar^2k^2}{2m}\,\psi_k(x)\,. \tag{13.7}$$

We require that $V(x) \to 0$ as $x \to \pm\infty$. So the wave function in the asymptotic regions is given by:

$$\psi_k(x) = \begin{cases} A\,e^{ikx} + B\,e^{-ikx}\,, & \text{as } x \to -\infty, \\ C\,e^{ikx} + D\,e^{-ikx}\,, & \text{as } x \to +\infty. \end{cases} \tag{13.8}$$

We define in and out coefficients by:

$$\Psi_{\rm in} = \begin{pmatrix} A \\ D \end{pmatrix}\,, \qquad \Psi_{\rm out} = \begin{pmatrix} C \\ B \end{pmatrix}\,. \tag{13.9}$$

The S-matrix is the connection between the in and out coefficients. That is, we define a 2×2 matrix S such that:

$$\Psi_{\rm out} = S\,\Psi_{\rm in}\,, \qquad S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}\,. \tag{13.10}$$

With our conventions, $S = 1$ when the potential vanishes. For particles incident from the left (negative x), $D = 0$ and the left-transmission and reflection coefficients are given by:

$$T_L = |S_{11}|^2\,, \qquad \text{and} \qquad R_L = |S_{21}|^2\,. \tag{13.11}$$

For particles incident from the right (positive x), $A = 0$ and the right-transmission and reflection coefficients are given by:

$$T_R = |S_{22}|^2\,, \qquad \text{and} \qquad R_R = |S_{12}|^2\,. \tag{13.12}$$

We can only find the S-matrix by a complete solution of Schrödinger's equation (13.7) for the particular potential $V(x)$. This is, in general, a difficult job. However, if the potential obeys certain properties, conservation laws and symmetry relations severely constrain the form of the S-matrix. We use these conservation laws and symmetry relations next to find a general form of S.

1. *Conservation of probability.* This means that current is conserved. The in and out currents are given by:

$$j_{\rm in} = \frac{\hbar k}{m}\,\left\{\,|A|^2 + |D|^2\,\right\}\,, \qquad \text{and} \qquad j_{\rm out} = \frac{\hbar k}{m}\,\left\{\,|C|^2 + |B|^2\,\right\}\,. \tag{13.13}$$

So since $j_{\rm in} = j_{\rm out}$, we have:

$$|A|^2 + |D|^2 = |C|^2 + |B|^2 \,, \tag{13.14}$$

which we can write as:

$$\Psi_{\rm in}^\dagger \, \Psi_{\rm in} = \Psi_{\rm out}^\dagger \, \Psi_{\rm out} = \Psi_{\rm in}^\dagger \, S^\dagger S \, \Psi_{\rm in} \,. \tag{13.15}$$

But this must be true for any in state, so S must be unitary:

$$S^\dagger S = S S^\dagger = 1 \,. \tag{13.16}$$

If S is given by:

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \,, \tag{13.17}$$

probability conservation means that:

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} S_{11}^* & S_{21}^* \\ S_{12}^* & S_{22}^* \end{pmatrix} = \begin{pmatrix} |S_{11}|^2 + |S_{12}|^2 & S_{11}S_{21}^* + S_{12}S_{22}^* \\ S_{21}S_{11}^* + S_{22}S_{12}^* & |S_{21}|^2 + |S_{22}|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{13.18}$$

2. *Time reversal.* As discussed in Section 13.3, time reversal invariance is a property of real potentials. It means that if $\psi_k(x)$ is a solution of Schrödinger's equation, then so is $\psi_k^*(x)$. We have explicitly used complex wave functions to describe waves moving in the left and right directions here, so we need to preserve this reality requirement with our asymptotic solutions. The complex conjugate of Eq. (13.19) is:

$$\psi_k^*(x) = \begin{cases} B^* \, e^{ikx} + A^* \, e^{-ikx} \,, & \text{as } x \to -\infty, \\ D^* \, e^{ikx} + C^* \, e^{-ikx} \,, & \text{as } x \to +\infty. \end{cases} \tag{13.19}$$

So now we find that

$$\begin{aligned} \Psi'_{\rm in} &= \begin{pmatrix} B^* \\ C^* \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} C^* \\ B^* \end{pmatrix} = Z \, \Psi_{\rm out}^* \,, \qquad \text{where} \qquad Z = Z^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \,, \\ \Psi'_{\rm out} &= \begin{pmatrix} D^* \\ A^* \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} A^* \\ D^* \end{pmatrix} = Z \, \Psi_{\rm in}^* \,. \end{aligned} \tag{13.20}$$

Now since

$$\Psi'_{\rm out} = S \, \Psi'_{\rm in} \,, \tag{13.21}$$

we find that

$$Z \, \Psi_{\rm in}^* = S \, Z \, \Psi_{\rm out}^* \,, \qquad \text{or} \qquad \Psi_{\rm out}^* = Z \, S^\dagger \, Z \, \Psi_{\rm in}^* \,, \tag{13.22}$$

which gives:

$$\Psi_{\rm out} = Z \, S^T \, Z \, \Psi_{\rm in} = S \, \Psi_{\rm in} \,, \tag{13.23}$$

so that $S = R \, S^T \, R$. If S is given by (13.17), this means that:

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} S_{22} & S_{12} \\ S_{21} & S_{11} \end{pmatrix} \,, \tag{13.24}$$

so that under time-reversal, $S_{11} = S_{22}$. This is the case for all *real* potentials.

3. *Parity.* Very often the potential is invariant under reversal of x. That is $V(-x) = V(x)$. Under parity, the wave function becomes:

$$\psi_k(x) = \begin{cases} D \, e^{ikx} + C \, e^{-ikx} \,, & \text{as } x \to -\infty, \\ B \, e^{ikx} + A \, e^{-ikx} \,, & \text{as } x \to +\infty. \end{cases} \tag{13.25}$$

So for this case,

$$\begin{aligned}\Psi''_{\rm in} &= \begin{pmatrix} D \\ A \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} A \\ D \end{pmatrix} = Z\,\Psi_{\rm in}\,, \\ \Psi''_{\rm out} &= \begin{pmatrix} B \\ C \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} C \\ B \end{pmatrix} = Z\,\Psi_{\rm out}\,,\end{aligned} \tag{13.26}$$

So since

$$\Psi''_{\rm out} = S\,\Psi''_{\rm in}\,, \tag{13.27}$$

we find that:

$$\Psi_{\rm out} = Z\,S\,Z\,\Psi_{\rm in} = S\,\Psi_{\rm in}\,, \qquad \text{which means that} \qquad S = Z\,S\,Z\,. \tag{13.28}$$

If S is given by (13.17), this means that:

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} S_{22} & S_{21} \\ S_{12} & S_{11} \end{pmatrix}\,. \tag{13.29}$$

So parity conservation requires that $S_{11} = S_{22}$ and $S_{12} = S_{21}$, which then means that $T_L = T_R$ and $R_L = R_R$, as expected from reflection invariance.

The S-matrix is a complex 2×2 matrix and so has a total of 8 real elements. Unitarity provides 4 independent equations, so this leaves 4 real elements. Time reversal provides only one additonal *independent* real equation, which then leave three independent elements for S. Parity then provides one more *independent* equation, which then leaves only two independent real elements to describe the S-matrix. After applying all these restrictions, we find the general form:

$$S = e^{i\phi} \begin{pmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{pmatrix}\,. \tag{13.30}$$

Exercise 25. Show that Eq. (13.30) satisfies unitarity, time reversal, and parity.

Any unitary matrix can be diagonalized by a unitary transformation. For our case, the eigenvalues of S are $e^{i(\phi\pm\theta)}$, and S is diagonalized by the matrix U, where

$$S_D = U^\dagger\,S\,U = \begin{pmatrix} e^{i(\phi+\theta)} & 0 \\ 0 & e^{i(\phi-\theta)} \end{pmatrix}\,, \qquad \text{where} \qquad U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\,. \tag{13.31}$$

The transmission of waves is simply described by S_D. If we put:

$$\Phi_{\rm in} = U\,\Psi_{\rm in} = \frac{1}{\sqrt{2}} \begin{pmatrix} A-D \\ A+D \end{pmatrix} \equiv \begin{pmatrix} a \\ d \end{pmatrix}\,, \qquad \text{and} \qquad \Phi_{\rm out} = U^\dagger\,\Psi_{\rm out} = \frac{1}{\sqrt{2}} \begin{pmatrix} C+B \\ C-B \end{pmatrix} \equiv \begin{pmatrix} c \\ b \end{pmatrix}\,, \tag{13.32}$$

then

$$\Phi_{\rm out} = S_D\,\Phi_{\rm in}\,, \qquad \text{or} \qquad c = e^{i(\phi+\theta)}\,a\,, \qquad b = e^{i(\phi-\theta)}\,d\,. \tag{13.33}$$

From the S matrix, we can find the transfer matrix M, which connects the coefficients on the left-hand side to coefficients on the right-hand side of the barrier. Let us define left L, right R vectors, and a transfer matrix M by:

$$L = \begin{pmatrix} A \\ B \end{pmatrix}\,, \qquad \text{and} \qquad R = \begin{pmatrix} C \\ D \end{pmatrix}\,, \qquad \text{with} \qquad R = M\,L\,. \tag{13.34}$$

Then, after some algebra, we find for symmetric potentials:

$$M = \begin{pmatrix} e^{-i\phi}\sec\theta & -i\tan\theta \\ +i\tan\theta & e^{+i\phi}\sec\theta \end{pmatrix}\,. \tag{13.35}$$

Note that $\det[\,M\,] = 1$, as required by current conservation.

Exercise 26. Find the S matrix for a square potential barrier $V(x)$ of the form:

$$V(x) = \begin{cases} 0\,, & \text{for } x < -a/2 \text{ and } x > +a/2, \\ V_0\,, & \text{for } -a/2 < x < +a/2, \end{cases} \tag{13.36}$$

for $E_k > V_0 > 0$. Show that your results agree with the general form of S given in Eq. (13.30). [See Merzbacher [**?**][p.93], but note that here we want to find S, not M.]

Exercise 27. Find the S matrix for the potential step $V(x)$, defined by

$$V(x) = \begin{cases} 0\,, & \text{for } x < 0, \\ V_0\,, & \text{for } x > 0. \end{cases}\,, \qquad \text{put} \qquad V_0 = \frac{\hbar^2\gamma^2}{2m}\,, \tag{13.37}$$

which defines γ. Consider the case when the kinetic energy of the particle for $x < 0$ is given by:

$$E = \frac{\hbar^2 k^2}{2m}\,, \tag{13.38}$$

and $E > V_0$. Put

$$\psi(x) = \begin{cases} A\,e^{+ikx} + B\,e^{-ikx}\,, & \text{for } x < 0, \\ C\,e^{+ik'x} + D\,e^{-ik'x}\,, & \text{for } x > 0. \end{cases}\,. \tag{13.39}$$

Find the relation between k and k'. If we define "in" and "out" states by:

$$\Psi_{\text{in}} = \begin{pmatrix} A \\ D \end{pmatrix}\,, \qquad \Psi_{\text{out}} = \begin{pmatrix} C \\ B \end{pmatrix}\,, \tag{13.40}$$

and, by applying the boundary conditions on the solutions at $x = 0$, find the 2×2 matrix S, defined by:

$$\Psi_{\text{out}} = S\,\Psi_{\text{in}}\,. \tag{13.41}$$

Show also that S obeys the probability conservation requirement:

$$S^\dagger\,K'\,S = K\,, \tag{13.42}$$

where K and K' are defined by:

$$K = \begin{pmatrix} k & 0 \\ 0 & k' \end{pmatrix}\,, \qquad K' = \begin{pmatrix} k' & 0 \\ 0 & k \end{pmatrix}\,. \tag{13.43}$$

Exercise 28. Suppose we want to use real functions on the left and right rather than complex ones. That is:

$$\psi_k(x) = \begin{cases} a\,\cos(kx) + b\,\sin(kx)\,, & \text{as } x \to -\infty, \\ c\,\cos(kx) + d\,\sin(kx)\,, & \text{as } x \to +\infty. \end{cases} \tag{13.44}$$

Using the results for the M matrix in (13.35), find the connection between c and d and a and b.

Exercise 29. Choose a gaussian barrier of the form:

$$V(x) = V_0\,e^{-x^2/L^2}\,, \tag{13.45}$$

with $E > V_0 > 0$, and, using the results of Exercise 28, and a numerical integrator (such as 4^{th} order Runge-Kutta), and find values for θ and ϕ. Choose convenient values for m, V_0, L, and E.

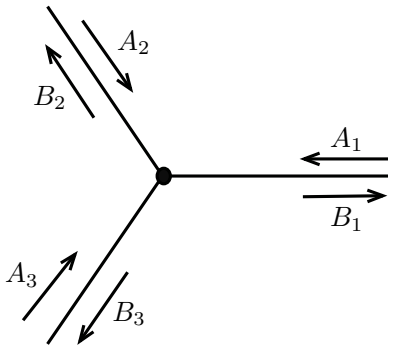

Figure 13.1: A junction with three legs.

Exercise 30. Consider a junction consisting of three one-dimensional legs, as shown in Fig. 13.1. Coefficients for the in and out wave functions for each leg are labeled as A_i and B_i for $i = 1, 2, 3$. We define in and out states as:

$$\Psi_{\rm in} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} , \qquad \Psi_{\rm out} = \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} , \tag{13.46}$$

and are connected by the S matrix: $\Psi_{\rm out} = S\,\Psi_{\rm in}$, where S is the 3×3 complex matrix:

$$S = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} , \tag{13.47}$$

and consists of 18 real elements.

1. Conservation of probability requires that S is unitary: $S^\dagger S = 1$. This requirement consists of 9 independent equations and reduces the number of independent elements in S to 9.

2. We also assume that the junction is symmetric with respect to each leg, so if we define a rotation matrix R by:

$$R = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} , \tag{13.48}$$

then

$$R\,\Psi_{\rm in} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = \begin{pmatrix} A_2 \\ A_3 \\ A_1 \end{pmatrix} = \Psi'_{\rm in} , \tag{13.49}$$

with a similar relation for the out states:

$$R\,\Psi_{\rm out} = \Psi'_{\rm out} . \tag{13.50}$$

Then since $\Psi'_{\rm out} = S\,\Psi'_{\rm in}$, we find:

$$R\,\Psi_{\rm out} = S\,R\,\Psi_{\rm in} , \qquad \text{or} \qquad \Psi_{\rm out} = R^{-1}\,S\,R\,\Psi_{\rm in} = S\,\Psi_{\rm in} . \tag{13.51}$$

So invariance under the first rotation requires that

$$S = R^{-1} S R . \tag{13.52}$$

Similarly for a second rotation, we find that

$$S = (R\,R)^{-1}\,S\,R\,R\,. \tag{13.53}$$

With these results, find the restrictions placed on the S-matrix by Eqs. (13.52) and (13.53). How many independent elements of S are left?

Solution: We first get R^{-1}. We find:

$$R = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}\,, \qquad \text{and} \qquad R^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\,. \tag{13.54}$$

Then

$$R^{-1}\,S\,R = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} S_{33} & S_{31} & S_{32} \\ S_{13} & S_{11} & S_{12} \\ S_{23} & S_{21} & S_{22} \end{pmatrix}\,. \tag{13.55}$$

So we conclude that:

$$S_{11} = S_{22} = S_{33} \equiv \alpha\,, \tag{13.56}$$

$$S_{31} = S_{12} = S_{23} \equiv \beta\,, \tag{13.57}$$

$$S_{32} = S_{13} = S_{21} \equiv \gamma\,. \tag{13.58}$$

A second rotation by R produces the same result, of course. So we conclude that S is of the form:

$$S = \begin{pmatrix} \alpha & \beta & \gamma \\ \gamma & \alpha & \beta \\ \beta & \gamma & \alpha \end{pmatrix}\,, \tag{13.59}$$

with α, β, and γ complex. Unitarity now requires:

$$\begin{aligned} S^\dagger S &= \begin{pmatrix} \alpha^* & \gamma^* & \beta^* \\ \beta^* & \alpha^* & \gamma^* \\ \gamma^* & \beta^* & \alpha^* \end{pmatrix} \begin{pmatrix} \alpha & \beta & \gamma \\ \gamma & \alpha & \beta \\ \beta & \gamma & \alpha \end{pmatrix} \\ &= \begin{pmatrix} |\alpha|^2 + |\beta|^2 + |\gamma|^2 & \alpha^*\beta + \gamma^*\alpha + \beta^*\gamma & \alpha^*\gamma + \gamma^*\beta + \beta^*\alpha \\ \beta^*\alpha + \alpha^*\gamma + \gamma^*\beta & |\alpha|^2 + |\beta|^2 + |\gamma|^2 & \beta^*\gamma + \alpha^*\beta + \gamma^*\alpha \\ \gamma^*\alpha + \beta^*\gamma + \alpha^*\beta & \gamma^*\beta + \beta^*\alpha + \alpha^*\gamma & |\alpha|^2 + |\beta|^2 + |\gamma|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{13.60}$$

There are only two independent equations here, which are:

$$|\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1\,, \tag{13.61}$$

$$\alpha\beta^* + \beta\gamma^* + \gamma\alpha^* = 0\,. \tag{13.62}$$

So we have left here a total of three complex numbers or six real numbers in the parameterization of S. So let us put:

$$\alpha = r_1\,e^{i\phi_1}\,, \qquad \beta = r_2\,e^{i\phi_2}\,, \qquad \gamma = r_3\,e^{i\phi_3}\,. \tag{13.63}$$

with r_1, r_2, and r_3 all real and non-negative. Then Eq. (13.61) requires that r_1, r_2, and r_3 are on the unit circle:

$$r_1^2 + r_2^2 + r_3^2 = 1\,. \tag{13.64}$$

Eq. (13.62) then gives:

$$r_1 r_2\,e^{i(\phi_1-\phi_2)} + r_2 r_3\,e^{i(\phi_2-\phi_3)} + r_3 r_1\,e^{i(\phi_3-\phi_1)} = 0\,. \tag{13.65}$$

Eq. (13.64) means that there are only two independent values of r, which reduces the number of parameters to five. However Eq. (13.65) is an additional complex equation, or two real equations, which would seem to reduce the number of independent parameters to three. It is not clear exactly how to pick them, however.

13.2.2 Wave packet propagation

In this section we look at wave packet propagation in one-dimension. The solution of Schrödinger's wave equation for a free particle is given by:

$$\psi(x,t) = \int_{-\infty}^{+\infty} \mathrm{d}k\, A_k\, e^{i\Phi_k(x,t)}\,, \qquad \text{where} \qquad \Phi_k(x,t) = kx - \omega_k t\,, \qquad \text{with} \qquad \omega_k = \frac{\hbar\, k^2}{2m}\,. \tag{13.66}$$

For a wave packet moving in the positive x-direction, we assume that A_k is centered about a value of $k = k_0$. So let us put $k' = k - k_0$, and expand the phase $\Phi_k(x,t)$ in a power series about k_0:

$$\begin{aligned} \Phi_k(x,t) &= \Phi_{k_0}(x,t) + \left.\frac{\partial \Phi_k(x,t)}{\partial k}\right|_{k_0} k' + \cdots\,, \\ &= k_0 x - \omega_0 t + (\,x - v_0 t\,)\, k' + \cdots\,, \end{aligned} \tag{13.67}$$

where $v_0 = \hbar k_0/m$ is the velocity of the center of the wave packet. So then keeping only the first two terms in the expansion (13.67) gives:

$$\psi(x,t) \approx e^{i(k_0 x - \omega_0 t)} \int_{-\infty}^{+\infty} \mathrm{d}k'\, A_{k_0+k'}\, e^{ik'(x-v_0 t)}\,. \tag{13.68}$$

At $t = 0$, we assume that the center of the wave packet is located at a position $x = x_0$, so that:

$$\psi(x,0) = e^{ik_0 x} \int_{-\infty}^{+\infty} \mathrm{d}k'\, A_{k_0+k'}\, e^{ik' x}\,. \tag{13.69}$$

Then Eq. (13.68) can be written as:

$$\psi(x,t) \approx e^{-i\omega_0 t}\, \psi(x - v_0 t, 0)\,. \tag{13.70}$$

That is, the probability of finding the particle at a point x at time t is given by:

$$P(x,t) = |\,\psi(x,t)\,|^2 \approx |\,\psi(x - v_0 t, 0)\,|^2 = P(x - v_0 t, 0)\,. \tag{13.71}$$

So since the center of the packet at $t = 0$ is located at $x = x_0$, the center of the packet moves according to the classical equation:

$$x = x_0 + v_0 t\,. \tag{13.72}$$

Our approximate result in Eq. (13.71) represents motion of the packet without change of shape. In reality, spreading of the wave packet takes place. In order to account for this spreading, we would have to include the second order term in the expansion of the phase in Eq. (13.67), which we ignore here.

13.2.3 Time delays for reflection by a potential step

In this section, we compute the time-delay for scattering from an potential step. Schrödinger's equation for this problem is:

$$\left\{ -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \right\} \psi(x,t) = i\hbar \frac{\partial \psi(x,t)}{\partial t}\,, \qquad \text{where} \qquad V(x) = \begin{cases} 0\,, & \text{for } x < 0, \\ V_0\,, & \text{for } x > 0. \end{cases} \tag{13.73}$$

We put

$$\psi(x,t) = \int_0^{k_{\max}} \mathrm{d}k\, \psi_k(x)\, \exp[\,-i\omega_k\, t\,]\,, \qquad \text{where} \qquad \omega_k = \frac{\hbar\, k^2}{2m}\,, \tag{13.74}$$

where $\psi_k(x)$ satisfies:

$$\left\{ -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x) \right\} \psi_k(x) = \frac{\hbar^2 k^2}{2m}\psi_k(x) \,. \tag{13.75}$$

So we put

$$\psi_k(x) = \begin{cases} A_k\, e^{ikx} + B_k\, e^{-ikx} \,, & \text{for } x < 0, \\ C_k\, e^{-\kappa x} \,, & \text{for } x > 0, \end{cases} \tag{13.76}$$

where

$$\kappa^2 = \gamma^2 - k^2 \,, \qquad \text{and we have put:} \qquad V_0 = \frac{\hbar^2\gamma^2}{2m} \,. \tag{13.77}$$

So we take $k_{\max} = \gamma$. The boundary conditions at $x = 0$ require that:

$$A_k + B_k = C_k \,, \qquad ik\,(\,A_k - B_k\,) = -\kappa\, C_k \,, \tag{13.78}$$

from which we find:

$$B_k = \frac{k - i\kappa}{k + i\kappa}\, A_k \,, \qquad C_k = \frac{2k}{k + i\kappa}\, A_k \,. \tag{13.79}$$

So let us put

$$k + i\kappa = \rho\, e^{i\phi_k} \,, \qquad \rho = \sqrt{k^2 + \kappa^2} = \gamma \,, \qquad \tan\phi_k = \frac{\kappa}{k} = \sqrt{\frac{\gamma^2}{k^2} - 1} \,, \tag{13.80}$$

with $0 < k < \gamma$. Putting these results into (13.79) gives

$$B_k = e^{-2i\phi_k}\, A_k \,, \qquad C_k = 2\, e^{-i\phi_k} \cos(\phi_k)\, A_k \,. \tag{13.81}$$

So Eq. (13.76) becomes:

$$\psi_k(x) = A_k \left\{ \left\{ e^{ikx} + e^{-i[kx + 2\phi_k]} \right\} \Theta(-x) + 2\, \cos(\phi_k)\, e^{-\kappa x - i\phi_k}\, \Theta(x) \right\} . \tag{13.82}$$

Substitution into Eq. (13.74) gives:

$$\psi(x,t) = \int_0^\gamma \mathrm{d}k\, A_k \left\{ \left\{ e^{i\Phi_k^{(1)}(x,t)} + e^{i\Phi_k^{(2)}(x,t)} \right\} \Theta(-x) + 2\, \cos(\phi_k)\, e^{\Phi_k^{(3)}(x,t)}\, \Theta(x) \right\} , \tag{13.83}$$

where the phases are given by:

$$\begin{aligned} \Phi_k^{(1)}(x,t) &= kx - \omega_k t \,, \\ \Phi_k^{(2)}(x,t) &= -kx - \omega_k t - 2\phi_k \,, \\ \Phi_k^{(3)}(x,t) &= -\kappa x - i\omega_k t - i\phi_k \,. \end{aligned} \tag{13.84}$$

The initial conditions are such that at $t = 0$, the center of the wave packet is located at a position $x = -L$ and moving towards positive x with an average velocity $v_0 = \hbar k_0/m$. That is A_k is centered about a positive value k_0, so let us put $k' = k - k_0$, and expand the phases about $k = k_0$. This gives:

$$\begin{aligned} \Phi_k^{(1)} &= k_0 x - \omega_0 t + (x - v_0 t)\, k' + \cdots \\ \Phi_k^{(2)} &= -k_0 x - \omega_0 t - 2\phi_0 - \left(x + v_0 t + 2\frac{\mathrm{d}\phi_k}{\mathrm{d}k} \right) k' + \cdots \\ &= -k_0 x - \omega_0 t - 2\phi_0 - \big(x + v_0(t - \tau) \big)\, k' + \cdots \,, \end{aligned} \tag{13.85}$$

where we have defined τ by:

$$\tau = -\frac{2}{v_0}\frac{\mathrm{d}\phi_k}{\mathrm{d}k} > 0 \,. \tag{13.86}$$

Substituting these results into (13.83) gives, for $x < 0$:

$$\begin{aligned}\psi(x,t) &\approx \int_{-\infty}^{\infty} \mathrm{d}k'\, A_{k_0+k'} \left\{ e^{i(k_0x-\omega_0t)+ik'(x-v_0t)} + e^{i(-k_0x-\omega_0t-2\phi_0)-ik'(x+v_0(t-\tau))} \right\} \\ &= e^{i(k_0x-\omega_0t)} \int_{-\infty}^{\infty} \mathrm{d}k'\, A_{k_0+k'}\, e^{ik'(x-v_0t)} + e^{i(-k_0x-\omega_0t-2\phi_0)} \int_{-\infty}^{\infty} \mathrm{d}k'\, A_{k_0+k'}\, e^{ik'(-x-v_0(t-\tau))} \,.\end{aligned} \tag{13.87}$$

At $t = 0$, the wave packet is represented by the first term in (13.87):

$$\psi(x,0) = e^{ik_0x} \int_{-\infty}^{\infty} \mathrm{d}k'\, A_{k_0+k'}\, e^{ik'x} \,. \tag{13.88}$$

So (13.87) becomes:

$$\psi(x,t) \approx e^{-i\omega_0t}\psi(x-v_0t,0) + e^{i(-2k_0x-\omega_0t-2\phi_0)}\psi(-x-v_0(t-\tau),0) \,. \tag{13.89}$$

The first term is a right-moving packet centered about $x = -L + v_0t$, and is the incident wave packet. For this term, x is *negative* for value of t between $0 < t < L/v_0$. The second term is a left-moving wave packet located at $x = L - v_0(t-\tau)$. For this term, x is *negative* for values of $t > \tau + L/v_0$. So we can interpret τ as the time that the particle spends *inside* the potential barrier.

13.3 Schrödinger's equation in two dimensions

Get this stuff from Tim Londergan's papers!

Exercise 31. Find the resonate energies for an electron confined to a two-dimensional circular ring bounded by $r = a$ and $r = b > a$.
*Solution:*In polar coordinates, the equation for the wave function $\psi(r,\theta)$ is:

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r}\frac{\partial}{\partial r}\,r\left(\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right]\psi(r,\theta) = E\,\psi(r,\theta) \,. \tag{13.90}$$

We want to find the eigenvalues of this equation such that $\psi(a,\theta) = \psi(b,\theta) = 0$, $a < b$. We first put $E = \hbar^2k^2/(2m)$, and find the equation:

$$\left[\frac{1}{r}\frac{\partial}{\partial r}\,r\left(\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2} + k^2\right]\psi(r,\theta) = 0 \,. \tag{13.91}$$

We separate variables by setting

$$\psi(r,\theta) = \psi_m(r)\, e^{im\theta} \,, \tag{13.92}$$

with m and integer, $-\infty \le m \le +\infty$, and get:

$$\left[\frac{1}{r}\frac{\partial}{\partial r}\,r\left(\frac{\partial}{\partial r}\right) - \frac{m^2}{r^2} + k^2\right]\psi_m(r) = 0 \,. \tag{13.93}$$

This is Bessel's equation, with the general solutions:

$$\psi_m(r) = A_m\, J_m(kr) + B_m\, N_m(kr) \,, \tag{13.94}$$

for $m \ge 0$, and which must vanish at $r = a$ and $r = b$. The first can be satisfied by taking $A_m = N_m(ka)$ and $B_m = -J_m(ka)$. Then we have:

$$\psi_m(r) = N_m(ka)\, J_m(kr) - J_m(ka)\, N_m(kr) \,. \tag{13.95}$$

This will vanish only if k is chosen to be one of the zeros of the equation:

$$N_m(k_{m,n}a)\,J_m(k_{m,n}b) - J_m(k_{m,n}a)\,N_m(k_{m,n}b) = 0\,, \tag{13.96}$$

where $k_{m,n}$ is the n^{th} zero of the m^{th} of Eq. (13.96). One needs to plot this function for typical values of a and b to see that there *are*, in fact, an infinite number of zeros of this function.

So eigenfunctions are given by:

$$\psi_{n,m}(r,\theta) = \{\, N_m(k_{n,m}a)\,J_m(k_{n,m}r) - J_m(k_{n,m}a)\,N_m(k_{n,m}r)\,\} \times \left\{\,\mathcal{N}_{n,m}\,e^{+im\theta} + \mathcal{N}_{n,-m}\,e^{-im\theta}\right\}\,, \tag{13.97}$$

for $m = 0, 1, 2, \ldots, +\infty$ and $n = 1, 2, 3, \ldots, +\infty$, and where $\mathcal{N}_{n,m}$ are arbitrary constants.

References

Chapter 14

The WKB approximation

14.1 Introduction

Sometimes the phase of the wave function in quantum mechanics is slowly varying. Under such circumstances, we might seek to find a non-perturbative expansion of the phase. The WKB approximation provides a systematic method of this expansion, and since it does not depend on the strength of the potential, it provides a useful way to understand the dynamics of the system. It can be applied to one dimensional wave mechanics only. We study this method in this chapter, and apply it to bound states and scattering problems. The key to applying this method is to find solutions to match the WKB-approximate solutions at the turning points of the potential. These are called the "connection formulas," and are discussed in Section 14.3 below.

14.2 Theory

We start with Schrödinger's equation in one-dimension:

$$\left\{ -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x) \right\} \psi(x) = E\,\psi(x)\,. \tag{14.1}$$

The WKB approximation is generated by considering a solution of the form:

$$\psi(x) = e^{iS(x)/\hbar}\,. \tag{14.2}$$

Then Schrödinger's equation becomes:

$$(\,S'(x)\,)^2 = p^2(x) + i\hbar\,S''(x)\,, \tag{14.3}$$

where

$$p^2(x) = 2m\,(E - V(x))\,. \tag{14.4}$$

If we ignore the second derivative term on the right-hand side of Eq. (14.3), we have approximately:

$$S'(x) = p(x) = \sqrt{2m\,(E - V(x))}\,, \tag{14.5}$$

or

$$S(x) = \int^x p(x)\,\mathrm{d}x\,. \tag{14.6}$$

Putting this solution back into the right-hand side of Eq. (14.3) gives the next order:

$$(\,S'(x)\,)^2 = p^2(x) + i\hbar\,p'(x)\,, \tag{14.7}$$

or

$$S'(x) = \sqrt{p^2(x) + i\hbar p'(x)} = p(x)\sqrt{1 + i\hbar p'(x)/p^2(x)} = p(x) + i\hbar\frac{p'(x)}{2p(x)} + \cdots , \tag{14.8}$$

or

$$\frac{\mathrm{d}S(x)}{\mathrm{d}x} = p(x) + i\hbar\frac{\mathrm{d}\ln\big[\sqrt{p(x)}\big]}{\mathrm{d}x} + \cdots , \tag{14.9}$$

so

$$S(x) = i\hbar\ln\big[\sqrt{p(x)}\big] + \int^x p(x)\,\mathrm{d}x\,. \tag{14.10}$$

To this order then, the WKB wave function is given by:

$$\psi(x) = \frac{1}{\sqrt{p(x)}}\exp\left\{i\int^x p(x)\,\mathrm{d}x/\hbar\right\}, \qquad \text{for } E > V(x). \tag{14.11}$$

In regions wher $E < V(x)$, we put $p(x) = i\tilde{p}(x)$, where $\tilde{p}(x) = \sqrt{2m(V(x) - E)}$, and solutions are given by:

$$\psi(x) = \frac{1}{\sqrt{\tilde{p}(x)}}\exp\left\{\int^x \tilde{p}(x)\,\mathrm{d}x/\hbar\right\}, \qquad \text{for } E < V(x). \tag{14.12}$$

The WKB approximation is generally carried out only to second order. The end points are fixed by boundary conditions, which will be explained in the next section.

14.3 Connection formulas

At the classical turning points where $p(x) = 0$, the WKB solutions blow up. So we will need to find exact solutions near turning points and match them with the WKB solutions we found in the last section.

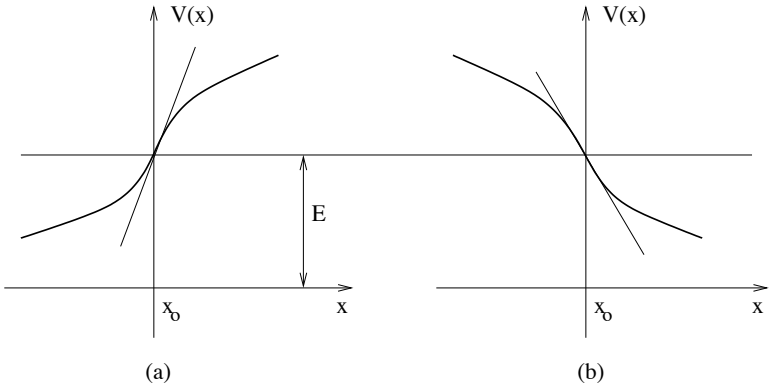

Figure 14.1: Two turning point situations.

14.3.1 Positive slope

We first examine the situation where the derivative of the potential at the turning point is *positive*, as shown in Fig. 14.1(a). The WKB solutions are given by:

$$\psi(x) = \begin{cases} \dfrac{A}{\sqrt{p(x)}}\exp[+i\int_x^{x_0} p(x)\mathrm{d}x/\hbar] + \dfrac{B}{\sqrt{p(x)}}\exp[-i\int_x^{x_0} p(x)\mathrm{d}x/\hbar]\,, & x < x_0, \\ \dfrac{C}{\sqrt{\tilde{p}(x)}}\exp[+\int_{x_0}^{x} \tilde{p}(x)\mathrm{d}x/\hbar] + \dfrac{D}{\sqrt{\tilde{p}(x)}}\exp[-\int_{x_0}^{x} \tilde{p}(x)\mathrm{d}x/\hbar]\,, & x > x_0. \end{cases} \tag{14.13}$$

We wish to find relations between the constants A, B, C, and D across the turning point region. We will do this by using an exact solution in the overlap region, assuming a linear potential in this region:

$$V(x) = V(x_0) + V'(x_0)\,(x - x_0) + \cdots\,, \tag{14.14}$$

with $E = V(x_0)$ and $V'(x_0) > 0$. Then the *exact* solution to Schrödinger's equation, to linear order, is given by:

$$\left\{ -\frac{\hbar^2}{2m}\,\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x_0) + V'(x_0)\,(x - x_0) \right\} \psi(x) = E\,\psi(x)\,, \tag{14.15}$$

or

$$\left\{ \frac{\mathrm{d}^2}{\mathrm{d}x^2} - \alpha^3\,(x - x_0) \right\} \psi(x) = 0\,. \tag{14.16}$$

where $\alpha^3 = 2m\,V'(x_0)/\hbar^2$. Setting $z = \alpha(x - x_0)$, we find the equation:

$$\left\{ \frac{\mathrm{d}^2}{\mathrm{d}z^2} - z \right\} \psi(z) = 0\,, \tag{14.17}$$

which has *Airy* functions as solutions. Airy functions are related to Bessel functions of order 1/3, and are throughly discussed by Abramowitz and Stegun[1, p. 446]. Two linearly independent Airy functions are written as $\mathrm{Ai}(z)$ and $\mathrm{Bi}(z)$. The asymptotic forms of these Airy functions are given by:

$$\mathrm{Ai}(z) \sim \begin{cases} \dfrac{1}{2\sqrt{\pi}\,z^{1/4}}\,e^{-2z^{3/2}/3}\,, & \text{for } z \to +\infty, \\[2ex] \dfrac{1}{\sqrt{\pi}(-z)^{1/4}}\,\sin\bigl[\,2(-z)^{3/2}/3 + \pi/4\,\bigr] & \text{for } z \to -\infty, \end{cases} \tag{14.18}$$

and

$$\mathrm{Bi}(z) \sim \begin{cases} \dfrac{1}{\sqrt{\pi}\,z^{1/4}}\,e^{+2z^{3/2}/3}\,, & \text{for } z \to +\infty, \\[2ex] \dfrac{1}{\sqrt{\pi}(-z)^{1/4}}\,\cos\bigl[\,2(-z)^{3/2}/3 + \pi/4\,\bigr] & \text{for } z \to -\infty. \end{cases} \tag{14.19}$$

A general solution of $\psi(x)$ in the vicinity of $x = x_0$ is given by the linear combination of the Airy functions:

$$\psi_p(x) = a\,\mathrm{Ai}[\alpha(x - x_0)] + b\,\mathrm{Bi}[\alpha(x - x_0)]\,, \tag{14.20}$$

with a and b constants to be fixed by matching with the WKB solutions. We call this wave function the *patching* wave function. The asymptotic forms for the patching wave function are given by:

$$\psi_p(x) \sim \begin{cases} \dfrac{1}{\sqrt{\pi}\,z^{1/4}} \left\{ b\,e^{+2z^{3/2}/3} + \dfrac{a}{2}\,e^{-2z^{3/2}/3} \right\}, & x < x_0, \\[2ex] \dfrac{1}{2i\sqrt{\pi}(-z)^{1/4}} \left\{ (a + ib)\,e^{+i[2(-z)^{3/2}/3 + \pi/4]} - (a - ib)\,e^{-i[2(-z)^{3/2}/3 + \pi/4]} \right\}, & x > x_0. \end{cases} \tag{14.21}$$

Here $z = \alpha(x - x_0)$. For the WKB solutions, in the patching region we have:

$$\begin{aligned} \tilde{p}(x) &= \sqrt{2m[V(x_0) + V'(x_0)(x - x_0) + \cdots - E]} = \hbar\alpha\,[\alpha(x - x_0)]^{1/2}\,, \\ \frac{1}{\hbar}\int_x^{x_0} \tilde{p}(x)\,\mathrm{d}x &= \alpha^{3/2}\int_{x_0}^{x} (x - x_0)^{1/2}\,\mathrm{d}x = 2\,[\alpha(x - x_0)]^{3/2}/3\,, \end{aligned} \tag{14.22}$$

for $x \gg x_0$ and

$$p(x) = \sqrt{2m[E - V(x_0) - V'(x_0)(x-x_0) + \cdots]} = \hbar\alpha\,[-\alpha(x-x_0)]^{1/2}\,,$$
$$\frac{1}{\hbar}\int_x^{x_0} p(x)\,\mathrm{d}x = \alpha^{3/2}\int_x^{x_0}(x_0-x)^{1/2}\,\mathrm{d}x = 2\,[-\alpha(x-x_0)]^{3/2}/3\,. \tag{14.23}$$

for $x \ll x_0$. So in the patching region, the WKB solutions are given by:

$$\psi(x) \sim \begin{cases} \dfrac{1}{\sqrt{\hbar\alpha}z^{1/4}}\left\{\, C\,e^{+2z^{3/2}/3} + D\,e^{-2z^{3/2}/3}\,\right\}, & \text{for } x \gg x_0, \\ \dfrac{1}{\sqrt{\hbar\alpha}(-z)^{1/4}}\left\{\, A\,e^{+i2(-z)^{3/2}/3} + B\,e^{-i2(-z)^{3/2}/3}\,\right\}, & \text{for } x \ll x_0, \end{cases} \tag{14.24}$$

where $z = \alpha(x - x_0)$. Comparing Eqs. (14.21) with (14.24), we find that

$$a = \sqrt{\frac{4\pi}{\hbar\alpha}}\,D\,, \qquad b = \sqrt{\frac{\pi}{\hbar\alpha}}\,C\,, \tag{14.25}$$

and

$$(a+ib)\,e^{+i\pi/4} = +i\sqrt{\frac{4\pi}{\hbar\alpha}}\,A\,, \qquad (a-ib)\,e^{-i\pi/4} = -i\sqrt{\frac{4\pi}{\hbar\alpha}}\,B\,, \tag{14.26}$$

from which we find the relations:

$$a = i\sqrt{\frac{\pi}{\hbar\alpha}}\,(A\,e^{-i\pi/4} - B\,e^{+i\pi/4}) = 2\sqrt{\frac{\pi}{\hbar\alpha}}\,D\,, \tag{14.27}$$

and

$$b = \sqrt{\frac{\pi}{\hbar\alpha}}\,(A\,e^{-i\pi/4} + B\,e^{+i\pi/4}) = \sqrt{\frac{\pi}{\hbar\alpha}}\,C\,. \tag{14.28}$$

Eliminating now the patching wave function constants a and b, we find the relations we seek between the WKB constants to the right and left of the turning point:

$$\frac{C}{2} = \frac{1}{2}\left[\,A\,e^{-i\pi/4} + B\,e^{+i\pi/4}\,\right],$$
$$D = \frac{i}{2}\left[\,A\,e^{-i\pi/4} - B\,e^{+i\pi/4}\,\right], \tag{14.29}$$

or

$$A = (\,C/2 - i\,D\,)\,e^{+i\pi/4}\,,$$
$$B = (\,C/2 + i\,D\,)\,e^{-i\pi/4}\,. \tag{14.30}$$

14.3.2 Negative slope

We now turn to the case when the derivative of the potential at the turning point is *negative*, as shown in Fig. 14.1(b). Here, we write the WKB solutions in the form:

$$\psi(x) = \begin{cases} \dfrac{A}{\sqrt{p(x)}}\,\exp[\,+i\int_{x_0}^{x} p(x)\mathrm{d}x/\hbar\,] + \dfrac{B}{\sqrt{p(x)}}\,\exp[\,-i\int_{x_0}^{x} p(x)\mathrm{d}x/\hbar\,], & x > x_0, \\ \dfrac{C}{\sqrt{\tilde{p}(x)}}\,\exp[\,+\int_{x}^{x_0} \tilde{p}(x)\mathrm{d}x/\hbar\,] + \dfrac{D}{\sqrt{\tilde{p}(x)}}\,\exp[\,-\int_{x}^{x_0} \tilde{p}(x)\mathrm{d}x/\hbar\,], & x < x_0. \end{cases} \tag{14.31}$$

For this case,

$$V(x) = V(x_0) + V'(x_0)\,(x - x_0) + \cdots\,, \tag{14.32}$$

with $E = V(x_0)$ and $V'(x_0) < 0$. The exact solutions are again Airy functions, but with negative arguement:

$$\psi_p(x) = a\,\mathrm{Ai}[-\alpha(x - x_0)] + b\,\mathrm{Bi}[-\alpha(x - x_0)]\,, \tag{14.33}$$

where α is now defined by: $\alpha = [\,-2mV'(x_0)/\hbar^2\,]^{1/3} > 0$. The asymptotic forms for the patching wave function are now given by:

$$\psi_p(x) \sim \begin{cases} \dfrac{1}{2i\sqrt{\pi}z^{1/4}}\left\{\,(a+ib)\,e^{+i[2z^{3/2}/3+\pi/4]} - (a-ib)\,e^{-i[2z^{3/2}/3+\pi/4]}\,\right\}, & x \gg x_0, \\ \dfrac{1}{\sqrt{\pi}(-z)^{1/4}}\left\{\,b\,e^{+2(-z)^{3/2}/3} + \dfrac{a}{2}\,e^{-2(-z)^{3/2}/3}\right\}, & x \ll x_0, \end{cases} \tag{14.34}$$

Here $z = \alpha(x - x_0)$.

For the WKB solutions, in the patching region we have:

$$\begin{aligned} p(x) &= \sqrt{2m[E - V(x_0) - V'(x_0)(x - x_0) + \cdots]} = \hbar\alpha\,[\alpha(x - x_0)]^{1/2}\,, \\ \frac{1}{\hbar}\int_{x_0}^{x} p(x)\,\mathrm{d}x &= \alpha^{3/2}\int_{x_0}^{x}(x - x_0)^{1/2}\,\mathrm{d}x = 2\,[\alpha(x - x_0)]^{3/2}/3\,, \end{aligned} \tag{14.35}$$

for $x \gg x_0$ and

$$\begin{aligned} \tilde{p}(x) &= \sqrt{2m[V(x_0) + V'(x_0)(x - x_0) + \cdots - E]} = \hbar\alpha\,[-\alpha(x - x_0)]^{1/2}\,, \\ \frac{1}{\hbar}\int_{x}^{x_0} \tilde{p}(x)\,\mathrm{d}x &= \alpha^{3/2}\int_{x}^{x_0}(x_0 - x)^{1/2}\,\mathrm{d}x = 2\,[-\alpha(x - x_0)]^{3/2}/3\,, \end{aligned} \tag{14.36}$$

for $x \ll x_0$. So in the patching region, the WKB solutions are given by:

$$\psi_(x) \sim \begin{cases} \dfrac{1}{\sqrt{\hbar\alpha}z^{1/4}}\left\{\,A\,e^{+i2z^{3/2}/3} + B\,e^{-i2z^{3/2}/3}\,\right\}, & x \gg x_0, \\ \left\{\,C\,e^{+2(-z)^{3/2}/3} + D\,e^{-2(-z)^{3/2}/3}\,\right\}, & x \ll x_0, \end{cases} \tag{14.37}$$

where $z = \alpha(x - x_0)$. Comparing Eqs. (14.34) with (14.37), we find that

$$(a+ib)\,e^{+i\pi/4} = +i\sqrt{\frac{4\pi}{\hbar\alpha}}\,A\,, \qquad (a-ib)\,e^{-i\pi/4} = -i\sqrt{\frac{4\pi}{\hbar\alpha}}\,B\,, \tag{14.38}$$

and

$$a = \sqrt{\frac{4\pi}{\hbar\alpha}}\,D\,, \qquad b = \sqrt{\frac{\pi}{\hbar\alpha}}\,C\,, \tag{14.39}$$

which are the *same* relations as Eqs. (14.25) and (14.26). So we find the same answers here as before:

$$a = i\sqrt{\frac{\pi}{\hbar\alpha}}\,(A\,e^{-i\pi/4} - B\,e^{+i\pi/4}) = 2\sqrt{\frac{\pi}{\hbar\alpha}}\,D\,, \tag{14.40}$$

and

$$b = \sqrt{\frac{\pi}{\hbar\alpha}}\,(A\,e^{-i\pi/4} + B\,e^{+i\pi/4}) = \sqrt{\frac{\pi}{\hbar\alpha}}\,C\,, \tag{14.41}$$

and eliminating the patching wave function constants a and b, we find the relations we seek between the WKB constants to the right and left of the turning point:

$$\begin{aligned} \frac{C}{2} &= \frac{1}{2}\left[\, A\, e^{-i\pi/4} + B\, e^{+i\pi/4}\,\right], \\ D &= \frac{i}{2}\left[\, A\, e^{-i\pi/4} - B\, e^{+i\pi/4}\,\right], \end{aligned} \tag{14.42}$$

or

$$\begin{aligned} A &= (\, C/2 - i\, D\,)\, e^{+i\pi/4}\,, \\ B &= (\, C/2 + i\, D\,)\, e^{-i\pi/4}\,. \end{aligned} \tag{14.43}$$

14.4 Examples

We apply our results for turning points to the two examples below.

14.4.1 Bound states

For a bound state situation, we consider a simple potential well shown in Fig. 14.2. Here we take the WKB

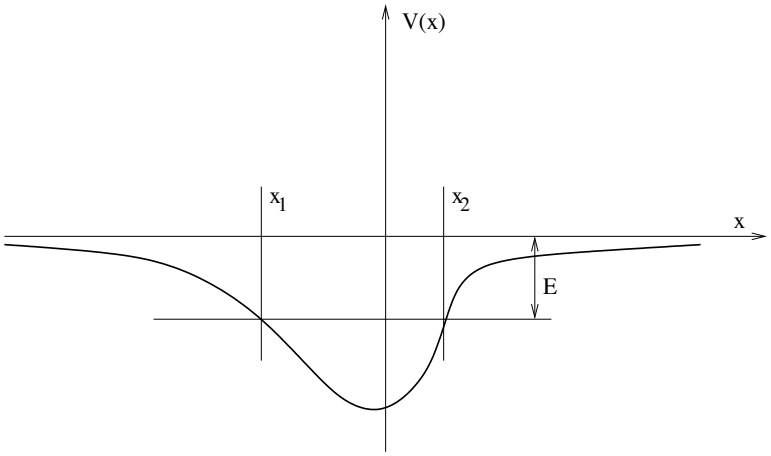

Figure 14.2: Potential well.

solutions to be of the form:

$$\psi(x) = \begin{cases} \dfrac{A}{\sqrt{\tilde{p}(x)}}\, \exp[\,-\int_x^{x_1} \tilde{p}(x)\mathrm{d}x/\hbar\,]\,, & x < x_1, \\ \dfrac{B}{\sqrt{p(x)}}\, \exp[\,+i\int_{x_1}^{x} p(x)\mathrm{d}x/\hbar\,] + \dfrac{C}{\sqrt{p(x)}}\, \exp[\,-i\int_{x_1}^{x} p(x)\mathrm{d}x/\hbar\,]\,, & x_1 < x < x_2, \\ \dfrac{D}{\sqrt{\tilde{p}(x)}}\, \exp[\,-\int_{x_2}^{x} \tilde{p}(x)\mathrm{d}x/\hbar\,]\,, & x_2 < x. \end{cases} \tag{14.44}$$

Applying Eq. (14.43) for the turning point at $x = x_1$, we find:

$$B = -iA\, e^{+i\pi/4}\,, \qquad C = +iA\, e^{-i\pi/4}\,. \tag{14.45}$$

In order to apply Eq. (14.30) for the turning point at $x = x_2$ we set:

$$\int_{x_1}^{x} p(x)\mathrm{d}x/\hbar = \Theta - \int_{x}^{x_2} p(x)\mathrm{d}x/\hbar\,, \qquad \text{where} \qquad \Theta = \int_{x_1}^{x_2} p(x)\mathrm{d}x/\hbar\,. \tag{14.46}$$

Then

$$\psi(x) = \frac{C\,e^{-i\Theta}}{\sqrt{p(x)}}\,\exp[+i\int_{x}^{x_2} p(x)\mathrm{d}x/\hbar] + \frac{B\,e^{+i\Theta}}{\sqrt{p(x)}}\,\exp[-i\int_{x}^{x_2} p(x)\mathrm{d}x/\hbar]\,, \tag{14.47}$$

for $x_1 < x < x_2$. This is now of the form required to apply Eq. (14.30) for the turning point at $x = x_2$. We find:

$$C\,e^{-i\Theta} = -i\,D\,e^{+i\pi/4}\,, \qquad B\,e^{+i\Theta} = +i\,D\,e^{-i\pi/4}\,. \tag{14.48}$$

So from Eqs. (14.45) and (14.48), we find:

$$D = -e^{+i(\Theta+\pi/2)}\,A = -e^{-i(\Theta+\pi/2)}\,A\,, \tag{14.49}$$

which requires that

$$\sin(\Theta + \pi/2) = 0\,, \tag{14.50}$$

so that:

$$\Theta = \int_{x_1}^{x_2} p(x)\mathrm{d}x/\hbar = (n + 1/2)\,\pi > 0\,, \tag{14.51}$$

where $n = 0, 1, 2, \ldots$. Note here that the values of n must be chosen so that Θ is non-negative. The WKB wave function is given by:

$$\psi_n(x) = A\begin{cases} \dfrac{1}{\sqrt{\tilde{p}(x)}}\,\exp[-\displaystyle\int_{x}^{x_1}\tilde{p}(x)\mathrm{d}x/\hbar]\,, & x < x_1, \\ \dfrac{2}{\sqrt{p(x)}}\,\sin\Big\{\displaystyle\int_{x_1}^{x} p(x)\mathrm{d}x/\hbar + \pi/4\Big\}\,, & x_1 < x < x_2, \\ \dfrac{(-)^n}{\sqrt{\tilde{p}(x)}}\,\exp[-\displaystyle\int_{x_2}^{x}\tilde{p}(x)\mathrm{d}x/\hbar]\,, & x_2 < x. \end{cases} \tag{14.52}$$

14.4.2 Tunneling

For a tunneling situation, we consider the simple potential barrier shown in Fig. 14.3. Here we take the WKB solutions to be of the form:

$$\psi(x) = \begin{cases} \dfrac{A}{\sqrt{p(x)}}\,\exp[-i\displaystyle\int_{x}^{x_1} p(x)\mathrm{d}x/\hbar] + \dfrac{B}{\sqrt{p(x)}}\,\exp[+i\displaystyle\int_{x}^{x_1} p(x)\mathrm{d}x/\hbar]\,, & x < x_1, \\ \dfrac{C}{\sqrt{\tilde{p}(x)}}\,\exp[+\displaystyle\int_{x_1}^{x}\tilde{p}(x)\mathrm{d}x/\hbar] + \dfrac{D}{\sqrt{\tilde{p}(x)}}\,\exp[-\displaystyle\int_{x_1}^{x}\tilde{p}(x)\mathrm{d}x/\hbar]\,, & x_1 < x < x_2, \\ \dfrac{F}{\sqrt{p(x)}}\,\exp[+i\displaystyle\int_{x_2}^{x} p(x)\mathrm{d}x/\hbar] + \dfrac{G}{\sqrt{p(x)}}\,\exp[-i\displaystyle\int_{x_2}^{x} p(x)\mathrm{d}x/\hbar]\,, & x_2 < x. \end{cases} \tag{14.53}$$

Here we must be careful to identify incoming and outgoing flux for the WKB solutions far from the scattering region. We first note that as $x \to \pm\infty$, $p(x) \sim p_0 = \sqrt{2mE}$, a constant. The probability flux is given by:

$$j(x) = \frac{\hbar}{2m\,i}\left\{\psi^*(x)\,\frac{\partial\psi(x)}{\partial x} + \frac{\partial\psi^*(x)}{\partial x}\,\psi(x)\right\}. \tag{14.54}$$

So as $x \to +\infty$, we find:

$$j_F(x) \sim +\frac{|F|^2}{m}\,, \qquad j_G(x) \sim -\frac{|G|^2}{m}\,, \tag{14.55}$$

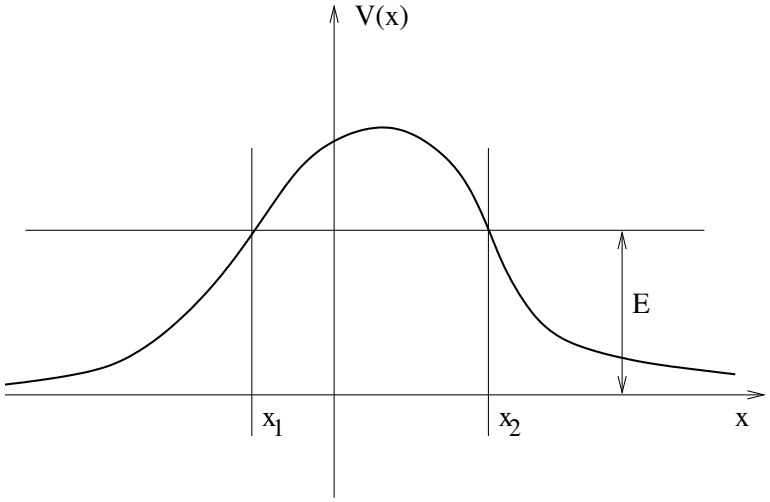

Figure 14.3: Potential barrier.

so F is the amplitude of the *outgoing* wave and G the amplitude for the *incoming* wave. For $x \to -\infty$, the situation is just reversed. Here, as $x \to -\infty$,

$$j_A(x) \sim +\frac{|A|^2}{m}\,, \qquad j_B(x) \sim -\frac{|B|^2}{m}\,, \tag{14.56}$$

so that A is the amplitude of the *outgoing* wave and B the amplitude for the *incoming* wave.

Applying now the connection formulas, Eqs. (14.30), for a turning point with positive slope at $x = x_1$, we find:

$$\begin{aligned} A &= (\,C/2 + i\,D\,)\,e^{-i\pi/4}\,, \\ B &= (\,C/2 - i\,D\,)\,e^{+i\pi/4}\,. \end{aligned} \tag{14.57}$$

This time in order to apply the connection forumulas Eqs. (14.42), for negative slope at $x = x_2$, we set:

$$\int_{x_1}^{x} \tilde{p}(x)\mathrm{d}x/\hbar = \tilde{\Theta} - \int_{x}^{x_2} \tilde{p}(x)\mathrm{d}x/\hbar\,, \qquad \text{where} \qquad \tilde{\Theta} = \int_{x_1}^{x_2} \tilde{p}(x)\,\mathrm{d}x\,/\hbar\,. \tag{14.58}$$

Then, the WKB solution for $x_1 < x < x_2$ becomes:

$$\psi(x) = \frac{D\,e^{-\tilde{\Theta}}}{\sqrt{\tilde{p}(x)}}\,\exp[+\int_{x}^{x_2} \tilde{p}(x)\mathrm{d}x/\hbar] + \frac{C\,e^{+\tilde{\Theta}}}{\sqrt{\tilde{p}(x)}}\,\exp[-\int_{x}^{x_2} \tilde{p}(x)\mathrm{d}x/\hbar]\,, \tag{14.59}$$

So now using Eq. (14.42), we find:

$$\begin{aligned} \frac{D}{2} &= \frac{1}{2}\left[\,F\,e^{-i\pi/4} + G\,e^{+i\pi/4}\,\right] e^{+\tilde{\Theta}} \\ C &= \frac{i}{2}\left[\,F\,e^{-i\pi/4} - G\,e^{+i\pi/4}\,\right] e^{-\tilde{\Theta}}\,. \end{aligned} \tag{14.60}$$

Combinining Eqs. (14.57) and (14.60), we find:

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \cosh\gamma & i\,\sinh\gamma \\ -i\,\sinh\gamma & \cosh\gamma \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix}\,, \tag{14.61}$$

where

$$\gamma = \ln[\,2\,e^{\tilde{\Theta}}\,] = \tilde{\Theta} + \ln 2\,, \tag{14.62}$$

so that

$$\sinh\gamma = \frac{1}{2}\left[\,2\,e^{\tilde{\Theta}} - \frac{1}{2\,e^{\tilde{\Theta}}}\,\right], \qquad \text{and} \qquad \cosh\gamma = \frac{1}{2}\left[\,2\,e^{\tilde{\Theta}} + \frac{1}{2\,e^{\tilde{\Theta}}}\,\right]. \tag{14.63}$$

Eq. (14.61) agrees with Liboff [2, p. 269] and Merzbacher [3][p. 126, Eq. (7.30)]. In and out coefficients and the S-matrix are defined by:

$$\Phi_{\rm in} = \begin{pmatrix} A \\ G \end{pmatrix}, \qquad \Phi_{\rm out} = \begin{pmatrix} F \\ B \end{pmatrix}, \qquad \Phi_{\rm out} = S\,\Phi_{\rm in}\,. \tag{14.64}$$

Rearranging Eq. (14.61), we find:

$$S = \begin{pmatrix} \mathrm{sech}\gamma & -i\tanh\gamma \\ -i\tanh\gamma & \mathrm{sech}\gamma \end{pmatrix}, \tag{14.65}$$

in agreement with Eq. (13.30) if we put $\mathrm{sech}\gamma = \cos\theta$ and then $\tanh\gamma = -\sin\theta$. So in the WKB approximation, the S-matrix is unitary. It also satisfies time reversal. However it also satisfies parity, even though we did not require that in our derivation of Eq. (14.65).

The right and left transmission and reflection coefficients are equal and are given by:

$$T_{\rm R} = T_{\rm L} = \mathrm{sech}^2\gamma\,, \qquad R_{\rm R} = R_{\rm L} = \tanh^2\gamma\,, \tag{14.66}$$

and add to one.

So the WKB approximation conserves probability.

In the limit $\tilde{\Theta} \to \infty$, we have:

$$\sinh\gamma = e^{\tilde{\Theta}}\left\{\,1 - \frac{1}{4}\,e^{-\tilde{\Theta}}\,\right\}, \qquad \cosh\gamma = e^{\tilde{\Theta}}\left\{\,1 + \frac{1}{4}\,e^{-\tilde{\Theta}}\,\right\}, \tag{14.67}$$

so that

$$T \sim e^{-2\tilde{\Theta}}\,, \qquad R \sim \left[\,\frac{1 - e^{-\tilde{\Theta}}/4}{1 + e^{-\tilde{\Theta}}/4}\,\right]^2 \sim 1 - e^{-2\tilde{\Theta}}\,. \tag{14.68}$$

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).

[2] R. L. Liboff, *Introductory Quantum Mechanics* (Addison-Wesley, awp:adr, 1997), third edition.

[3] E. Merzbacher, *Quantum Mechanics* (John Wiley & Sons, New York, NY, 1970), second edition.

Chapter 15

Spin systems

Spin-1/2 systems are particularly important. Electrons, protons, and neutrons all have spin 1/2. The behavior of these particles in various physical situations is important to understand because of the applications of these properties to useful devices, such as atomic clocks, electron and proton spin resonance, and microwave devices. These are all quantum devices that work according to quantum mechanics. Spin-1/2 systems also provide a means of analyzing *any* two-level system, and provide quantum solutions to such systems. For example, optical pumping of two-level systems can by analyzed by spinors.

15.1 Magnetic moments

The magnetic moment of a spin-1/2 particle is given by:

$$\boldsymbol{\mu} = \frac{q\lambda}{mc}\,\mathbf{S} = \frac{q\lambda\,\hbar}{2mc}\,\boldsymbol{\sigma}\,, \qquad \text{with} \qquad \mathbf{S} = \frac{\hbar}{2}\,\boldsymbol{\sigma}\,, \tag{15.1}$$

where σ_i are the Pauli matrices, defined below. For electrons, $q = -e$, $\lambda = 1$, and $m = m_e$ is the electron mass. For protons, $q = +e$, $\lambda = 1.397$, and $m = M_p$. For the neutron, $q = -e$, $\lambda = 0.957$, with $m = M_p$.

15.2 Pauli matrices

The Pauli matrices are defined by:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\,, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,, \qquad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\,.$$

The σ-matrices are all Hermitian and traceless. They obey the algebra,

$$\begin{aligned} [\,\sigma_i, \sigma_j\,] &= \sigma_i\sigma_j - \sigma_j\sigma_i = 2i\,\epsilon_{ijk}\sigma_k\,, \\ \{\,\sigma_i, \sigma_j\,\} &= \sigma_i\sigma_j + \sigma_j\sigma_i = 2\,\delta_{ij} I\,, \end{aligned} \tag{15.2}$$

from which we find:

$$\sigma_i\sigma_j = \delta_{ij} I + i\epsilon_{ijk}\sigma_k\,. \tag{15.3}$$

We also note that $\sigma_2\,\sigma_i\,\sigma_2 = -\sigma_i^*$. If $\mathbf{a}$ and $\mathbf{b}$ are vectors, then multiplying Eq. (15.3) by a_i and b_j gives:

$$(\mathbf{a}\cdot\boldsymbol{\sigma})\,(\mathbf{b}\cdot\boldsymbol{\sigma}) = (\mathbf{a}\cdot\mathbf{b})\,I + i(\mathbf{a}\times\mathbf{b})\cdot\boldsymbol{\sigma}\,. \tag{15.4}$$

Next, we establish trace formulas. From Eq. (15.3), we find that:

$$\mathrm{Tr}[\,\sigma_i\sigma_j\,] = 2\,\delta_{ij}\,, \tag{15.5}$$

from which we find:

$$\begin{aligned} \mathrm{Tr}[\,\mathbf{a}\cdot\boldsymbol{\sigma}\,] &= 0\,, \\ \mathrm{Tr}[\,(\mathbf{a}\cdot\boldsymbol{\sigma})\,(\mathbf{b}\cdot\boldsymbol{\sigma})\,] &= 2\,(\mathbf{a}\cdot\mathbf{b})\,, \\ \mathrm{Tr}[\,(\mathbf{a}\cdot\boldsymbol{\sigma})\,(\mathbf{b}\cdot\boldsymbol{\sigma})\,(\mathbf{c}\cdot\boldsymbol{\sigma})\,] &= 2i\,\mathbf{a}\cdot(\,\mathbf{b}\times\mathbf{c}\,)\,, \qquad \text{etc.} \end{aligned} \tag{15.6}$$

Any 2×2 matrix A can be written as:

$$A = \frac{1}{2}\,[\,a_0 + \mathbf{a}\cdot\boldsymbol{\sigma}\,]\,, \tag{15.7}$$

where, from Eq. (15.6), we find:

$$a_0 = \mathrm{Tr}[\,A\,]\,, \qquad \mathbf{a} = \mathrm{Tr}[\,\boldsymbol{\sigma}A\,]\,. \tag{15.8}$$

If A is Hermitian, then a_0 and $\mathbf{a}$ must be real.

15.2.1 The eigenvalue problem

In the next theorem, we solve the eigenvalue problem for the operator $\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}$.

Theorem 34. *The eigenvalue problem for the operator* $\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}$*:*

$$(\,\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}\,)\,\chi_\lambda(\hat{\mathbf{r}}) = \lambda\,\chi_\lambda(\hat{\mathbf{r}})\,, \tag{15.9}$$

where $\hat{\mathbf{r}}$ *is a unit vector given by:*

$$\begin{aligned} \hat{\mathbf{r}} &= x\,\hat{\mathbf{e}}_x + y\,\hat{\mathbf{e}}_y + z\,\hat{\mathbf{e}}_z\,, \qquad \textit{with} \qquad x^2+y^2+z^2 = 1\,, \\ &= \sin\theta\cos\phi\,\hat{\mathbf{e}}_x + \sin\theta\sin\phi\,\hat{\mathbf{e}}_y + \cos\theta\,\hat{\mathbf{e}}_z\,, \end{aligned} \tag{15.10}$$

has solutions with eigenvalues $\lambda = \pm 1$*, and eigenvectors* $\chi_\pm(\hat{\mathbf{r}})$*, given by:*

$$\begin{aligned} \chi_+(\hat{\mathbf{r}}) &= \begin{pmatrix} e^{-i\phi/2}\cos(\theta/2) \\ e^{+i\phi/2}\sin(\theta/2) \end{pmatrix} = \frac{e^{-i\phi/2}}{2\cos(\theta/2)}\begin{pmatrix} 1+x_3 \\ x_1+ix_2 \end{pmatrix} = \frac{e^{-i\phi/2}}{2\sin(\theta/2)}\begin{pmatrix} x_1-ix_2 \\ 1-x_3 \end{pmatrix}\,, \\ \chi_-(\hat{\mathbf{r}}) &= \begin{pmatrix} -e^{-i\phi/2}\sin(\theta/2) \\ e^{+i\phi/2}\cos(\theta/2) \end{pmatrix} = \frac{e^{+i\phi/2}}{2\cos(\theta/2)}\begin{pmatrix} -x_1+ix_2 \\ 1+x_3 \end{pmatrix} = \frac{e^{-i\phi/2}}{2\sin(\theta/2)}\begin{pmatrix} x_3-1 \\ x_1+ix_2 \end{pmatrix}\,. \end{aligned} \tag{15.11}$$

Proof. We note that:

$$\hat{\mathbf{r}}\cdot\boldsymbol{\sigma} = \begin{pmatrix} z & x-iy \\ x+iy & -z \end{pmatrix} = \begin{pmatrix} \cos\theta & e^{-i\phi}\sin\theta \\ e^{i\phi}\sin\theta & -\cos\theta \end{pmatrix}\,, \tag{15.12}$$

from which we can easily find eigenvalues and eigenvectors. The rest of the proof is straightforward, and we leave it for an exercise for the reader. □

Exercise 32. Find the eigenvalues and eigenvectors given in Theorem 34.

The unitary transformation which brings the matrix $(\hat{\mathbf{x}}\cdot\boldsymbol{\sigma})$ to diagonal form is then given by *row's* made up of the complex conjugate of the two eigenvalues:

$$D(\hat{\mathbf{r}}) = \begin{pmatrix} e^{+i\phi/2}\cos(\theta/2) & e^{-i\phi/2}\sin(\theta/2) \\ -e^{+i\phi/2}\sin(\theta/2) & e^{-i\phi/2}\cos(\theta/2) \end{pmatrix}\,. \tag{15.13}$$

Then:

$$D(\hat{\mathbf{r}})\,(\hat{\mathbf{r}}\cdot\boldsymbol{\sigma})\,D^\dagger(\hat{\mathbf{r}}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \hat{\mathbf{e}}_z\cdot\boldsymbol{\sigma}\,. \tag{15.14}$$

Exercise 33. Show that

$$D(\theta,\phi) = e^{i\hat{\mathbf{n}}(\phi)\cdot\boldsymbol{\sigma}\,\theta/2}\,, \qquad \text{where} \qquad \hat{\mathbf{n}}(\phi) = -\cos\phi\,\hat{\mathbf{e}}_x + \sin\phi\,\hat{\mathbf{e}}_y\,. \tag{15.15}$$

Prove also Eq. (15.14) by direct multiplication of the matrices. Draw a picture of a coordinate system for the eigenvalue problem, showing the vectors $\hat{\mathbf{r}}$ and $\hat{\mathbf{n}}$, and the angles θ and ϕ.

Exercise 34. Show also that the $D(\theta,\phi)$, given in Eq. (15.13) above, is related to the $D^{(1/2)}$ matrix defined by the Eular angles in Eq. (21.158), by the equation:

$$D(\theta,\phi) = D^{(1/2)}(0,\theta,\phi)\,. \tag{15.16}$$

Show in a diagram how these two coordinates are related to the Eular angles.

Definition 29. Projection operators are defined by:

$$\begin{aligned} P_+(\theta,\phi) &= \chi_+(\theta,\phi)\,\chi_+^\dagger(\theta,\phi) = \frac{1}{2}\,(\,1+\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}\,) = \begin{pmatrix} \cos^2(\theta/2) & e^{-i\phi}\sin(\theta)/2 \\ e^{i\phi}\sin(\theta)/2 & \sin^2(\theta/2)\end{pmatrix}\,, \\ P_-(\theta,\phi) &= \chi_-(\theta,\phi)\,\chi_-^\dagger(\theta,\phi) = \frac{1}{2}\,(\,1-\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}\,) = \begin{pmatrix} \sin^2(\theta/2) & -e^{-i\phi}\sin(\theta)/2 \\ -e^{i\phi}\sin(\theta)/2 & \cos^2(\theta/2)\end{pmatrix}\,. \end{aligned} \tag{15.17}$$

We note that $P_+(\theta,\phi) + P_-(\theta,\phi) = 1$.

Remark 25. We can think of $\chi_+(\theta,\phi)$ as representing the eigenstate of spin up in the $\hat{\mathbf{r}}$ direction, and $P_+(\theta,\phi)$ as the density matrix which describes a particle with spin up in the $\hat{\mathbf{r}}$ direction. *Any* eigenvector can be written as a linear combination of spin up and spin down eigenvectors with respect to *any* axis.

Exercise 35. Show that $P_\pm(\theta,\phi)$ project from any arbitrary spinor χ an eigenstate of $\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}$ with eigenvalue ± 1.

Exercise 36. Show that: $\mathrm{Tr}[\,P(\theta,\phi)\,] = 1$ and $\mathrm{Tr}[\,\boldsymbol{\sigma}P(\theta,\phi)\,] = \hat{\mathbf{r}}$.

15.3 Spin precession in a magnetic field

We consider a particle with magnetic moment $\boldsymbol{\mu}$ and spin-1/2 in a constant magnetic field of magnitude $\mathbf{B}_0$. The Hamiltonian is given by:

$$H = -\boldsymbol{\mu}\cdot\mathbf{B}_0 = -\frac{q\lambda\,\hbar}{2mc}\,\boldsymbol{\sigma}\cdot\mathbf{B}_0 = \frac{\hbar}{2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\omega}_0\,, \qquad \text{where} \qquad \boldsymbol{\omega}_0 = \gamma\,\mathbf{B}_0\,, \qquad \text{with} \qquad \gamma = -\frac{q\lambda}{mc}\,. \tag{15.18}$$

The commutation relations for spin at time t are:

$$[\,S_i(t), S_j(t)\,] = i\hbar\,\epsilon_{ijk}\,S_k(t)\,, \qquad \text{or} \qquad [\,\sigma_i(t),\sigma_j(t)\,] = 2i\,\epsilon_{ijk}\,\sigma_k(t)\,. \tag{15.19}$$

(i) In the Heisenberg picture, we have

$$\dot{\sigma}_i(t) = \frac{1}{i\hbar}\,[\,\sigma_i(t), H(t)\,] = \frac{1}{2i}\,[\,\sigma_i(t),\sigma_j(t)\omega_{0\,j}\,] = \epsilon_{ijk}\,\omega_{0\,j}\,\sigma_k(t)\,, \tag{15.20}$$

which we can write in a vector notation as:

$$\frac{\mathrm{d}\boldsymbol{\sigma}(t)}{\mathrm{d}t} = \boldsymbol{\omega}_0\times\boldsymbol{\sigma}(t)\,. \tag{15.21}$$

This equation represents precession of the operator $\boldsymbol{\sigma}(t)$ about the direction $\hat{\boldsymbol{\omega}}_0 = \boldsymbol{\omega}_0/\omega_0$. The solution is:

$$\boldsymbol{\sigma}(t) = (\,\boldsymbol{\sigma}(0)\cdot\hat{\boldsymbol{\omega}}_0\,)\,\hat{\boldsymbol{\omega}}_0 + (\,\hat{\boldsymbol{\omega}}_0\times\boldsymbol{\sigma}(0)\,)\times\hat{\boldsymbol{\omega}}_0\,\cos(\omega_0 t) + (\,\hat{\boldsymbol{\omega}}_0\times\boldsymbol{\sigma}(0)\,)\,\sin(\omega_0 t)\,. \tag{15.22}$$

Let us set $\hat{\mathbf{p}}(t) = \langle\, \boldsymbol{\sigma}(t)\,\rangle$ so that the average value of the spin is $\langle\, \mathbf{S}(t)\,\rangle = \hbar\,\mathbf{p}(t)/2$. Then we find:

$$\hat{\mathbf{p}}(t) = (\,\hat{\mathbf{p}}(0)\cdot\hat{\boldsymbol{\omega}}_0\,)\,\hat{\boldsymbol{\omega}}_0 + (\,\hat{\boldsymbol{\omega}}_0\times\hat{\mathbf{p}}(0)\,)\times\hat{\boldsymbol{\omega}}_0\,\cos(\omega_0 t) + (\,\hat{\boldsymbol{\omega}}_0\times\hat{\mathbf{p}}(0)\,)\,\sin(\omega_0 t)\,. \tag{15.23}$$

Now for a spin-1/2 system, there are no higher moments since $\langle\, \sigma_i^2(t)\,\rangle = 1$, and since the correlation coefficients, given by $\langle\, \sigma_i(t)\sigma_j(t)\,\rangle = i\epsilon_{ijk}\,\langle\, \sigma_k(t)\,\rangle = i\epsilon_{ijk}\,p_k(t)$, are related to average values of the spin, it turns out that the average value of the spin, $\hat{\mathbf{p}}(t)$ *completely describes the state of the system.* We can understand this result by considering the general form of the density matrix. In the Schrödinger representation, the density matrix is defined by:

$$\rho(t) := |\,\psi(t)\,\rangle\langle\,\psi(t)\,| = \frac{1}{2}\,(\,1+\hat{\mathbf{p}}(t)\cdot\boldsymbol{\sigma}\,)\,. \tag{15.24}$$

The density matrix satisfies:

$$\begin{aligned} \mathrm{Tr}[\,\rho(t)\,] &= \langle\,\psi(t)\,|\,\psi(t)\,\rangle = \langle\,\psi(0)\,|\,\psi(0)\,\rangle = 1\,,\\ \mathrm{Tr}[\,\boldsymbol{\sigma}\,\rho(t)\,] &= \langle\,\psi(t)\,|\,\boldsymbol{\sigma}\,|\,\psi(t)\,\rangle = \langle\,\boldsymbol{\sigma}(t)\,\rangle = \hat{\mathbf{p}}(t)\,. \end{aligned} \tag{15.25}$$

We also note that the density matrix is **idempotent**: $\rho(t)\rho(t) = \rho(t)$. So

$$\frac{1}{2}\,(\,1+\hat{\mathbf{p}}(t)\cdot\boldsymbol{\sigma}\,)\,\frac{1}{2}\,(\,1+\hat{\mathbf{p}}(t)\cdot\boldsymbol{\sigma}\,) = \frac{1}{4}\,(\,1+\hat{\mathbf{p}}(t)\cdot\hat{\mathbf{p}}(t)+2\,\hat{\mathbf{p}}(t)\cdot\boldsymbol{\sigma}\,) \equiv \frac{1}{2}\,(\,1+\hat{\mathbf{p}}(t)\cdot\boldsymbol{\sigma}\,)\,, \tag{15.26}$$

so $\hat{\mathbf{p}}(t)\cdot\hat{\mathbf{p}}(t) = 1$. That is $\hat{\mathbf{p}}(t)$ is a unit vector for all t. $\hat{\mathbf{p}}(t)$ is called the **polarization vector** for the spin-1/2 system.

(ii) In the Schrödinger picture, we want to solve Schrödinger's equation:

$$H\,|\,\psi(t)\,\rangle = \frac{\hbar}{2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\omega}_0\,|\,\psi(t)\,\rangle = i\hbar\,\frac{\partial|\,\psi(t)\,\rangle}{\partial t}\,. \tag{15.27}$$

We solve this problem by putting

$$|\,\psi(t)\,\rangle = \sum_i c_i\,e^{-iE_i t/\hbar}\,|\,\psi_i\,\rangle\,, \tag{15.28}$$

where $|\,\psi_i\,\rangle$ and E_i are eigenvectors and eigenvalues of the equation:

$$\frac{\hbar}{2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\omega}_0\,|\,\psi_i\,\rangle = E_i\,|\,\psi_i\,\rangle\,. \tag{15.29}$$

Putting $E_i = \hbar\,\omega_i/2$, the eigenvalue equation becomes:

$$\boldsymbol{\sigma}\cdot\boldsymbol{\omega}_0\,|\,\psi_i\,\rangle = \omega_i\,|\,\psi_i\,\rangle\,. \tag{15.30}$$

Solutions exist if:

$$\begin{vmatrix} \omega_{0\,z}-\omega & \omega_{0\,x}-i\omega_{0\,y}\\ \omega_{0\,x}+i\omega_{0\,y} & -\omega_{0\,z}-\omega \end{vmatrix} = \omega^2-(\omega_{0\,x}^2+\omega_{0\,y}^2+\omega_{0\,z}^2) = (\omega-\omega_0)(\omega+\omega_0) = 0\,. \tag{15.31}$$

So $\omega = \pm\omega_0$. From Theorem 34, the eigenvectors are given by:

$$|\,\boldsymbol{\omega}_0,+\,\rangle = \begin{pmatrix} e^{-i\alpha/2}\,\cos(\beta/2)\\ e^{+i\alpha/2}\,\sin(\beta/2)\end{pmatrix}\,,\qquad \text{and}\qquad |\,\boldsymbol{\omega}_0,-\,\rangle = \begin{pmatrix} -e^{-i\alpha/2}\,\sin(\beta/2)\\ e^{+i\alpha/2}\,\cos(\beta/2)\end{pmatrix}\,, \tag{15.32}$$

where (α,β) are the azimuthal and polar angles of the vector $\boldsymbol{\omega}_0$ in an arbitrary coordinate system. However the eigenvectors are much simpler if we choose this arbitrary coordinate system so that $\boldsymbol{\omega}_0$ is in the z-direction. Then $\alpha = \beta = 0$, and the eigenvectors become:

$$|\,\boldsymbol{\omega}_0,+\,\rangle = \begin{pmatrix}1\\0\end{pmatrix}\,,\qquad \text{and}\qquad |\,\boldsymbol{\omega}_0,-\,\rangle = \begin{pmatrix}0\\1\end{pmatrix}\,. \tag{15.33}$$

Then the general solution of Schrödinger's equation is:

$$|\,\psi(t)\,\rangle = c_+\, e^{-i\omega_0 t/2}\,|\,\boldsymbol{\omega}_0,+\,\rangle + c_-\, e^{+i\omega_0 t/2}\,|\,\boldsymbol{\omega}_0,-\,\rangle = \begin{pmatrix} c_+\, e^{-i\omega_0 t/2} \\ c_-\, e^{+i\omega_0 t/2}\end{pmatrix} \tag{15.34}$$

At $t = 0$, the spin state is an eigenvector with eigenvalue $+\hbar/2$ pointing in a direction specified by the polar angles (ϕ, θ), which is given by:

$$|\,\psi(0)\,\rangle = \begin{pmatrix} e^{-i\phi/2}\,\cos(\theta/2) \\ e^{+i\phi/2}\,\sin(\theta/2)\end{pmatrix} = \begin{pmatrix} c_+ \\ c_-\end{pmatrix}\,. \tag{15.35}$$

So the solution to Schrödinger's equation is:

$$|\,\psi(t)\,\rangle = \begin{pmatrix} e^{-i(\,\phi+\omega_0 t\,)/2}\,\cos(\theta/2) \\ e^{+i(\,\phi+\omega_0 t\,)/2}\,\sin(\theta/2)\end{pmatrix}\,. \tag{15.36}$$

The density matrix is then given by:

$$\begin{aligned}\rho(t) = |\,\psi(t)\,\rangle\langle\,\psi(t)\,| &= \begin{pmatrix} e^{-i(\,\phi+\omega_0 t\,)/2}\,\cos(\theta/2) \\ e^{+i(\,\phi+\omega_0 t\,)/2}\,\sin(\theta/2)\end{pmatrix}\Big(e^{+i(\,\phi+\omega_0 t\,)/2}\,\cos(\theta/2)\,,\quad e^{-i(\,\phi+\omega_0 t\,)/2}\,\sin(\theta/2)\Big) \\ &= \begin{pmatrix} \sin^2(\theta/2) & e^{+i(\,\phi+\omega_0 t\,)}\,\sin(\theta/2)\cos(\theta/2) \\ e^{-i(\,\phi+\omega_0 t\,)}\,\sin(\theta/2)\cos(\theta/2) & \cos^2(\theta/2)\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1+\cos(\theta) & e^{+i(\,\phi+\omega_0 t\,)}\,\sin(\theta) \\ e^{-i(\,\phi+\omega_0 t\,)}\,\sin(\theta) & 1-\cos(\theta)\end{pmatrix} = \frac{1}{2}\,(\,1+\hat{\mathbf{p}}(t)\cdot\boldsymbol{\sigma}\,)\,,\end{aligned} \tag{15.37}$$

where $\hat{\mathbf{p}}(t)$ is a unit vector given by:

$$\hat{\mathbf{p}}(t) = \sin(\theta)\,\cos(\phi+\omega_0 t)\,\hat{\mathbf{e}}_x + \sin(\theta)\,\sin(\phi+\omega_0 t)\,\hat{\mathbf{e}}_y + \cos(\theta)\,\hat{\mathbf{e}}_z\,, \tag{15.38}$$

which represents precession of the polarization vector $\hat{\mathbf{p}}(t)$ about the z-axis by an amount ω_0, and in agreement with the Heisenberg result, Eq. (15.23), for our special coordinate system.

In this section, we discuss the dynamics of a free spin-1/2 proton in a magnetic field $\mathbf{B}(t)$. The Hamiltonian for this system is given by:

$$H(t) = -\boldsymbol{\mu}\cdot\mathbf{B}(t)\,,\qquad \text{where}\qquad \boldsymbol{\mu} = \frac{\hbar\lambda}{2m}\,\boldsymbol{\sigma}\,, \tag{15.39}$$

where $\lambda = +1.123$ and m is the mass of the proton. Schrödinger's equation is given by:

$$H\,\chi(t) = i\hbar\,\frac{\mathrm{d}\chi(t)}{\mathrm{d}t}\,. \tag{15.40}$$

15.4 Driven spin system

In this example, we add a time-dependent external magnetic field $\mathbf{B}_1(t)$, perpendicular to $\mathbf{B}_0$, to the spin-1/2 system of example **??**. The Hamiltonian then takes the form:

$$H = -\frac{q\lambda\,\hbar}{2mc}\,\boldsymbol{\sigma}\cdot(\,\mathbf{B}_0+\mathbf{B}_1(t)\,) = \frac{\hbar}{2}\,\boldsymbol{\sigma}\cdot(\,\boldsymbol{\omega}_0+\boldsymbol{\omega}_1(t)\,)\,, \tag{15.41}$$

where

$$\boldsymbol{\omega}_0 = \gamma\,\mathbf{B}_0\,,\qquad \text{and}\qquad \boldsymbol{\omega}_1(t) = \gamma\,\mathbf{B}_1(t)\,. \tag{15.42}$$

So we put:

$$H_0 = \frac{\hbar}{2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\omega}_0\,, \qquad \text{and} \qquad H_1(t) = \frac{\hbar}{2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\omega}_1(t)\,. \tag{15.43}$$

So we find:

$$H_1'(t) = \frac{\hbar}{2}\,e^{+i\boldsymbol{\sigma}\cdot\boldsymbol{\omega}_0 t/2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\omega}_1(t)\,e^{-i\boldsymbol{\sigma}\cdot\boldsymbol{\omega}_0 t/2} \tag{15.44}$$

Let us fix the coordinate system so that $\boldsymbol{\omega}_0 = \omega_0\hat{\mathbf{e}}_z$. Let us also consider the case when $\boldsymbol{\omega}_1(t)$ rotates uniformly about the z-axis:

$$\boldsymbol{\omega}_1(t) = \omega_1\left[\,\sin(\beta)\,\cos(\gamma t)\,\hat{\mathbf{e}}_x + \sin(\beta)\,\sin(\gamma t)\,\hat{\mathbf{e}}_y\, + \cos(\beta)\,\hat{\mathbf{e}}_z\,\right]. \tag{15.45}$$

From the appendix, we find that:

$$\begin{aligned}
e^{+i\sigma_z\omega_0 t/2}\,\sigma_x\,e^{-i\sigma_z\omega_0 t/2} &= \sigma_x\,\cos(\omega_0 t) - \sigma_y\,\sin(\omega_0 t)\,,\\
e^{+i\sigma_z\omega_0 t/2}\,\sigma_y\,e^{-i\sigma_z\omega_0 t/2} &= \sigma_x\,\sin(\omega_0 t) + \sigma_y\,\cos(\omega_0 t)\,,\\
e^{+i\sigma_z\omega_0 t/2}\,\sigma_z\,e^{-i\sigma_z\omega_0 t/2} &= \sigma_z\,.
\end{aligned} \tag{15.46}$$

So $H_1'(t)$ becomes:

$$\begin{aligned}
H_1'(t) &= \frac{\hbar}{2}\left\{\left[\,\sigma_x\,\cos(\omega_0 t) - \sigma_y\,\sin(\omega_0 t)\,\right]\omega_{1\,x}(t) + \left[\,\sigma_x\,\sin(\omega_0 t) + \sigma_y\,\cos(\omega_0 t)\,\right]\omega_{1\,y}(t) + \sigma_z\,\omega_{1\,z}(t)\,\right\}\\
&= \frac{\hbar}{2}\left\{\,\sigma_x\left[\,\omega_{1\,x}(t)\,\cos(\omega_0 t) + \omega_{1\,y}(t)\,\sin(\omega_0 t)\,\right]\right.\\
&\qquad\left. + \sigma_y\left[\,\omega_{1\,y}(t)\,\cos(\omega_0 t) - \omega_{1\,x}(t)\,\sin(\omega_0 t)\,\right] + \sigma_z\,\omega_{1\,z}(t)\,\right\}\\
&= \frac{\hbar\omega_1}{2}\left\{\,\sigma_x\,\sin(\beta)\left[\,\cos(\gamma t)\,\cos(\omega_0 t) + \sin(\gamma t)\,\sin(\omega_0 t)\,\right]\right.\\
&\qquad\left. + \sigma_y\,\sin(\beta)\left[\,\sin(\gamma t)\,\cos(\omega_0 t) - \cos(\gamma t)\,\sin(\omega_0 t)\,\right] + \sigma_z\,\cos(\beta)\,\right\}\\
&= \frac{\hbar\omega_1}{2}\left\{\,\sigma_x\,\sin(\beta)\,\cos(\omega t) + \sigma_y\,\sin(\beta)\,\sin(\omega t) + \sigma_z\,\cos(\beta)\,\right\},
\end{aligned} \tag{15.47}$$

where $\omega = \gamma - \omega_0$. It is now useful to transform to a coordinate system rotating about the z-axis by an amount ω. So let

$$\begin{aligned}
\omega_x &= \omega_x'\,\cos(\omega t) - \omega_y'\,\sin(\omega t)\,,\\
\omega_y &= \omega_x'\,\sin(\omega t) + \omega_y'\,\cos(\omega t)\,,\\
\omega_z &= \omega_z'\,.
\end{aligned} \tag{15.48}$$

Then (15.47) becomes:

$$H'(t) = \frac{\hbar\omega_1}{2}\left\{\,\sin(\beta)\,\sigma_x' + \cos\beta\,\sigma_z'\,\right\} = \frac{\hbar}{2}\,\mathbf{B}'\cdot\boldsymbol{\sigma}'\,, \tag{15.49}$$

where

$$\mathbf{B}' = \omega_1\left(\,\sin(\beta)\,\hat{\mathbf{e}}_x' + \cos(\beta)\,\hat{\mathbf{e}}_z'\,\right) = \omega_1\,\hat{\mathbf{n}}'\,, \tag{15.50}$$

where

$$\hat{\mathbf{n}}' = \sin(\beta)\,\hat{\mathbf{e}}_x' + \cos(\beta)\,\hat{\mathbf{e}}_z'\,, \qquad \text{and} \qquad |\,\mathbf{B}'\,| = \omega_1\,. \tag{15.51}$$

In this coordinate system, $\mathbf{B}'$ is *independent* explicitly of time. So this is now the same problem as what we solved in Example **??** for a constant magnetic field only now, the magnetic field is pointed in the $\hat{\mathbf{n}}'$ direction in the *rotating system* and has magnitude ω_1. This is illustrated in Fig. 15.1. So the solution in the rotating coordinate system is precession about $\hat{\mathbf{n}}'$ by an amount ω_1. So the polarization vector in this rotating coordinate system is given by Eq. (15.23), evaluated in the rotating system:

$$\hat{\mathbf{p}}'(t) = (\,\hat{\mathbf{p}}'(0)\cdot\hat{\mathbf{n}}'\,)\,\hat{\mathbf{n}}' + (\,\hat{\mathbf{n}}'\times\hat{\mathbf{p}}'(0)\,)\times\hat{\mathbf{n}}'\,\cos(\omega_1 t) + (\,\hat{\mathbf{n}}'\times\hat{\mathbf{p}}'(0)\,)\,\sin(\omega_1 t)\,. \tag{15.52}$$

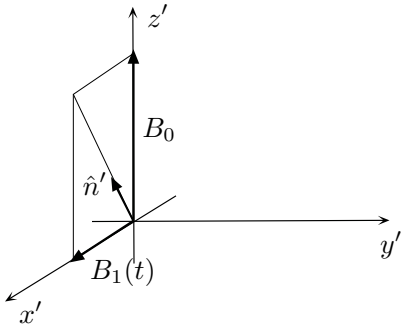

Figure 15.1: Spin precession in the rotating coordinate system.

Recall that β is the angle between $\mathbf{B}_0$ and $\mathbf{B}_1(t)$. If $\beta = \pi/2$ so that $\hat{\mathbf{n}}' = \hat{\mathbf{e}}'_x$ and if at $t = 0$ the polarization points in the negative z'-direction $\hat{\mathbf{p}}'(0) = -\hat{\mathbf{e}}'_z$, as shown in Fig. 15.1, the polarization vector traces a circle in the $x'z'$-plane of the rotating coordinate system, with the magnitude of the polarization going the full range from $-1 \leq p(t) \leq +1$. The $x'y'x'$ coordinate system rotates with respect to the laboratory fixed system with a frequency $\omega = \gamma - \omega_0$, so if the system is tuned so that $\omega = 0$, the $x'y'z'$-system is also fixed in the laboratory. This results in a resonance state where the spin system absorbs electromagnetic energy and reradiates this energy from the $\mathbf{B}_1(t)$ field. If ω is some other value, the plane of the polarization vector rotates, either clockwise or counter clockwise depending on the sign of ω. If $\beta \neq \pi/2$, the polarization vector never completely flips over.

15.5 Spin decay: T_1 and T_2

Here we discuss decay of spin systems caused by interactions with magnetic fields produced by other atoms.

15.6 The Ising model

Here we discuss the Ising model.

15.7 Heisenberg models

Here we discuss the Heisenberg xx and xy spin models.

References

Chapter 16

The harmonic oscillator

Harmonic oscillation occurs in many branches of quantum physics and is an important motion to study in detail. In this chapter we discuss quantization of the classical system, the eigenvalue problem, coherent and squeezed states, and the forced oscillator.

We also discuss the fermi oscilator and its relation to supersymmetry.

16.1 The Lagrangian

The classical Lagrangian for a particle subject to a harmonic restoring force in one-dimension is given by:

$$L(q,\dot{q}) = \frac{1}{2}\, m\,(\,\dot{q}^2 - \omega_0^2\, q^2\,)\,. \tag{16.1}$$

It is useful to first remove the units from this problem. Let us define the oscillator length parameter b by:

$$b = \sqrt{\frac{\hbar}{m\omega_0}}\,, \tag{16.2}$$

and put $\bar{q} = q/b$. If we set $\tau = \omega_0 t$ and put

$$\dot{q} \equiv \frac{\mathrm{d}q}{\mathrm{d}t} = b\,\omega_0\,\frac{\mathrm{d}\bar{q}}{\mathrm{d}\tau} \equiv b\,\omega_0\,\bar{q}'\,, \tag{16.3}$$

then $\bar{q}'$ has no units. The Lagrangian then becomes:

$$L(q,\dot{q}) = \frac{\hbar\,\omega_0}{2}\,(\,\bar{q}'^{\,2} - \bar{q}^2\,) \equiv \hbar\,\omega_0\,\bar{L}(\bar{q},\bar{q}')\,, \tag{16.4}$$

so that $\bar{L}(\bar{q},\bar{q}')$ has no units. So let us just revert back to using the unbarred coordinate system and dots instead of primes and assume throughout this section that we have scaled the units as above. We then have the Lagrangian:

$$L(q,\dot{q}) = \frac{1}{2}\,(\,\dot{q}^2 - q^2\,)\,, \qquad p = \frac{\partial L(q,\dot{q})}{\partial \dot{q}} = \dot{q}\,. \tag{16.5}$$

The Hamiltonian is:

$$H(q,p) = \frac{1}{2}\,(\,p^2 + q^2\,)\,, \qquad \text{and} \qquad \{\,q,p\,\} = 1\,. \tag{16.6}$$

Our scaling is equivalent to setting:

$$m = \omega_0 = \hbar = 1\,. \tag{16.7}$$

To recover ordinary units, one can perform the following replacements:

$$q \mapsto q/b\,, \qquad p \mapsto b\,p/\hbar\,, \qquad t \mapsto \omega_0 t\,, \qquad H \mapsto \hbar\omega\,H\,. \tag{16.8}$$

Here we have introduced a scale, $\hbar$, into the *classical* system in anticipation of canonical quantization. So we now map $q \mapsto Q$ and $p \mapsto P$ to Hermitian operators in Hilbert space, with the result:

$$H(Q,P) = \frac{1}{2}\,(\,P^2 + Q^2\,)\,, \qquad \text{and} \qquad [\,Q,P\,] = i\,. \tag{16.9}$$

The time development operator is

$$U(t) = e^{-iHt}\,. \tag{16.10}$$

The Heisenberg equations of motion are:

$$\begin{aligned} \dot{Q} &= [\,Q,H\,]/i = P\,, \\ \dot{P} &= [\,P,H\,]/i = -Q\,, \end{aligned} \tag{16.11}$$

so that:

$$\ddot{Q} + Q = 0\,, \qquad \ddot{P} + P = 0\,. \tag{16.12}$$

These equations have solutions:

$$\begin{aligned} Q(t) &= U^\dagger(t)\,Q\,U(t) = Q\,\cos t + P\,\sin t\,, \\ P(t) &= U^\dagger(t)\,P\,U(t) = P\,\cos t - Q\,\sin t\,. \end{aligned} \tag{16.13}$$

Here Q and P are time-independent Hermitian operators. We can put:[1]

$$\begin{aligned} Q(t) &= \frac{1}{\sqrt{2}}\,\big(\,A(t) + A^\dagger(t)\,\big)\,, \qquad & A(t) &= \frac{1}{\sqrt{2}}\,\big(\,Q(t) + iP(t)\,\big)\,, \\ P(t) &= \frac{1}{i\sqrt{2}}\,\big(\,A(t) - A^\dagger(t)\,\big)\,, \qquad & A^\dagger(t) &= \frac{1}{\sqrt{2}}\,\big(\,Q(t) - iP(t)\,\big)\,, \end{aligned} \tag{16.15}$$

so that $[\,A,A^\dagger\,] = 1$. The equations of motion for $A(t)$ and $A^\dagger(t)$ are:

$$\dot{A} = [\,A,H\,]/i = -i\,A\,, \qquad \dot{A}^\dagger = [\,A^\dagger,H\,]/i = +i\,A^\dagger\,, \tag{16.16}$$

which have the solutions:

$$A(t) = U^\dagger(t)\,A\,U(t) = A\,e^{-it}\,, \qquad A^\dagger(t) = U^\dagger(t)\,A^\dagger\,U(t) = A^\dagger\,e^{+it}\,. \tag{16.17}$$

We will use these results in this chapter.

16.2 Energy eigenvalue and eigenvectors

It is useful to obtain solutions to the energy eigenvalue problem. This is defined by the equation:

$$H\,|\,n\,\rangle = E_n\,|\,n\,\rangle\,, \qquad H = \frac{1}{2}\,(\,P^2 + Q^2\,)\,, \tag{16.18}$$

[1] In ordinary units,

$$\begin{aligned} Q(t) &= \frac{b}{\sqrt{2}}\,\big(\,A(t) + A^\dagger(t)\,\big)\,, \qquad & A(t) &= \frac{1}{\sqrt{2}}\,\big(\,Q(t)/b + ib\,P(t)/\hbar\,\big)\,, \\ P(t) &= \frac{\hbar}{i\sqrt{2}b}\,\big(\,A(t) - A^\dagger(t)\,\big)\,, \qquad & A^\dagger(t) &= \frac{1}{\sqrt{2}}\,\big(\,Q(t)/b - ib\,P(t)/\hbar\,\big)\,, \end{aligned} \tag{16.14}$$

$A(t)$ and $A^\dagger(t)$ have no units.

where the operators Q and P are at $t = 0$. The eigenvalue problem is easily solved by putting at $t = 0$,

$$Q = \left(A + A^\dagger \right)/\sqrt{2}\,, \qquad P = \left(A - A^\dagger \right)/i\sqrt{2}\,, \tag{16.19}$$

so that $[\, A, A^\dagger \,] = 1$, and

$$H = \frac{1}{2}\left(A^\dagger A + A A^\dagger \right) = N + 1/2\,, \qquad N = A^\dagger A\,. \tag{16.20}$$

The eigenvalue problem for H then reduces to that for N, where we can use the results of the next theorem:

Theorem 35 (The number operator)**.** *The eigenvalues and eigenvectors of the number operator N:*

$$N\,|\,n\,\rangle = n\,|\,n\,\rangle\,, \qquad N = A^\dagger A\,, \qquad [\, A, A^\dagger \,] = 1\,, \tag{16.21}$$

is given by $n = 0, 1, 2, \ldots$, with

$$A\,|\,n\,\rangle = \sqrt{n}\,|\,n-1\,\rangle\,, \qquad A^\dagger\,|\,n\,\rangle = \sqrt{n+1}\,|\,n+1\,\rangle\,, \tag{16.22}$$

and

$$|\,n\,\rangle = \frac{[\,A^\dagger\,]^n}{\sqrt{n!}}\,|\,0\,\rangle\,. \tag{16.23}$$

Proof. We start by noting that N is a positive definite operator and therefore has a lower bound:

$$\langle\, n\,|\,N\,|\,n\,\rangle = \langle\, n\,|A^\dagger A|\,n\,\rangle = \langle\, A\,n\,|\,A\,n\,\rangle = n\,\langle\, n\,|\,n\,\rangle\,, \tag{16.24}$$

so

$$n = \frac{|\,A|\,n\,\rangle|^2}{|\,|\,n\,\rangle\,|^2} \geq 0\,. \tag{16.25}$$

We also note that:

$$[\, A, N \,] = A\,, \qquad [\, A^\dagger, N \,] = -A^\dagger\,, \tag{16.26}$$

So

$$\begin{aligned} N\left\{\, A\,|\,n\,\rangle \,\right\} &= \left\{\, A\,N - [\, A, N \,] \,\right\}|\,n\,\rangle \quad = (n-1)\left\{\, A\,|\,n\,\rangle \,\right\}, \\ N\left\{\, A^\dagger\,|\,n\,\rangle \,\right\} &= \left\{\, A^\dagger\,N - [\, A^\dagger, N \,] \,\right\}|\,n\,\rangle = (n+1)\left\{\, A^\dagger\,|\,n\,\rangle \,\right\}. \end{aligned}$$

from which we find:

$$A\,|\,n\,\rangle = c_n\,|\,n-1\,\rangle\,, \qquad A^\dagger\,|\,n\,\rangle = d_n\,|\,n+1\,\rangle\,,$$

and so if the states $|\,n\,\rangle$ are normalized to one for all n, we find:

$$\begin{aligned} n &= \langle\, n\,|\,A^\dagger\,A\,|\,n\,\rangle \;=\; |\,c_n\,|^2\,\langle\, n-1\,|\,n-1\,\rangle \;=\; |\,c_n\,|^2 \\ n+1 &= \langle\, n\,|\,A\,A^\dagger\,|\,n\,\rangle \;=\; |\,d_n\,|^2\,\langle\, n+1\,|\,n+1\,\rangle \;=\; |\,d_n\,|^2\,, \end{aligned}$$

choosing the phases to be one, we find $c_n = \sqrt{n}$ and $d_n = \sqrt{n+1}$. This gives the results:

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

This means that there is a lowest state, call it $|\,n_0\,\rangle$ such that

$$A\,|\,n_0\,\rangle = 0\,.$$

But this means that $N\,|\,n_0\,\rangle = A^\dagger\,A\,|\,n_0\,\rangle = n_0|\,n_0\,\rangle = 0$, so that $n_0 = 0$. Therefore the eigenvalues are: $n = 0, 1, 2, \ldots$. The eigenvectors $|\,n\,\rangle$ are then obtained by successive application of $A^\dagger$ on the ground state $|\,0\,\rangle$. The result of this, by induction, is:

$$|\,n\,\rangle = \frac{[\,A^\dagger\,]^n}{\sqrt{n!}}\,|\,0\,\rangle\,.$$

This completes the proof. □

To find the wave functions in the coordinate representation, we start by noting that $A\,|\,0\,\rangle = 0$ defined the ground state. Now since $A = (Q + iP)/\sqrt{2}$,

$$\langle\, q\,|\,A\,|\,0\,\rangle = \frac{1}{\sqrt{2}}\,\langle\, q\,|(\,iP + Q\,)\,|\,0\,\rangle = \frac{1}{\sqrt{2}}\left\{\,\frac{\mathrm{d}}{\mathrm{d}q} + q\,\right\}\psi_0(q) = 0\,,$$

where $\psi_0(q) = \langle\, q\,|\,0\,\rangle$. The normalized solution is given by:

$$\psi_0(q) = \frac{1}{\pi^{1/4}}\,e^{-\frac{1}{2}q^2}\,, \qquad \int_{-\infty}^{\infty}|\psi_0(q)|^2\,\mathrm{d}q = 1\,. \tag{16.27}$$

Recall that q is written in units of the oscillator length b. For the states $\phi_n(q)$ with $n > 0$, we apply the $A^\dagger$ operator in coordinate space on $\psi_0(q)$ n times. This gives:

$$\begin{aligned}\psi_n(q) = \langle\, q\,|\,n\,\rangle &= \frac{1}{\sqrt{n!}}\langle\, q\,|\,\big[\,A^\dagger\,\big]^n\,|\,0\,\rangle = \frac{1}{\sqrt{2\,n!}}\langle\, q\,|\big[\,-iP + Q\,\big]^n|\,0\,\rangle \\ &= \frac{(-1)^n}{2^{n/2}\sqrt{n!}}\left[\,\frac{\mathrm{d}}{\mathrm{d}q} - q\,\right]^n\psi_0(x) = \frac{(-1)^n}{\pi^{1/4}\,2^{n/2}\sqrt{n!}}\left[\,\frac{\mathrm{d}}{\mathrm{d}q} - q\,\right]^n e^{-\frac{1}{2}q^2} \\ &= \frac{1}{\pi^{1/4}\,2^{n/2}\sqrt{n!}}\,H_n(q)\,e^{-\frac{1}{2}q^2}\,,\end{aligned}$$

where we have used the definition of Hermite polynomials:

$$H_n(q) = (-1)^n\,e^{\frac{1}{2}q^2}\left[\,\frac{\mathrm{d}}{\mathrm{d}q} - q\,\right]^n e^{-\frac{1}{2}q^2} = (-1)^n\left[\,e^{q^2}\,\frac{\mathrm{d}^n}{\mathrm{d}q^n}\,e^{-q^2}\,\right].$$

All wave functions are normalized with respect to x:

$$\int_{-\infty}^{\infty}|\psi_n(q)|^2\,\mathrm{d}q = 1\,.$$

In the momentum representation,

$$\langle\, p\,|\,A\,|\,0\,\rangle = \frac{1}{\sqrt{2}}\,\langle\, p\,|(\,iP + Q\,)\,|\,0\,\rangle = \frac{i}{\sqrt{2}}\left\{\,\frac{\mathrm{d}}{\mathrm{d}p} + p\,\right\}\tilde{\psi}_0(p) = 0\,,$$

the normalized solution of which is given by:

$$\tilde{\psi}_0(p) = \frac{\sqrt{2\pi}}{\pi^{1/4}}\,e^{-\frac{1}{2}p^2}\,, \qquad \int_{-\infty}^{\infty}|\tilde{\psi}_0(p)|^2\,\frac{\mathrm{d}p}{2\pi} = 1\,.$$

Exercise 37. Show that the harmonic oscillator wave functions in the momentum representation is given by:

$$\tilde{\psi}_n(p) = \frac{\sqrt{2\pi}}{\pi^{1/4}\,2^{n/2}\sqrt{n!}}\,H_n(p)\,e^{-\frac{1}{2}p^2}\,. \tag{16.28}$$

16.3 Other forms of the Lagrangian

Starting with the Lagrangian given in Eq. (16.5), let us define $x = q$ and $y = \dot{q}$, so that in terms of these variables, we can write the Lagrangian, which we here call L_1, in a number of different ways:

$$\begin{aligned}L_1(x,y;\dot{x},\dot{y}) &= \frac{1}{2}\,(\,y^2 - x^2\,) \\ &= y\dot{x} - \frac{1}{2}\,(\,x^2 + y^2\,) \\ &= \frac{1}{2}\,(\,y\dot{x} - x\dot{y} - x^2 - y^2\,) + \frac{1}{2}\,\frac{\mathrm{d}(xy)}{\mathrm{d}t}\,,\end{aligned} \tag{16.29}$$

But since a total derivative cannot change variation of the action, the Lagrangian

$$L_2(x,y;\dot{x},\dot{y}) = \frac{1}{2}\left(\, y\dot{x} - x\dot{y} - x^2 - y^2 \,\right) \tag{16.30}$$

must lead to the same equations of motion that we started with. Indeed, using Lagrangian (16.30) we find:

$$\begin{aligned} p_x &= \frac{\partial L_2}{\partial \dot{x}} = +\frac{y}{2}\,, \qquad \frac{\partial L_2}{\partial x} = -\frac{\dot{y}}{2} - x\,, \\ p_y &= \frac{\partial L_2}{\partial \dot{y}} = -\frac{x}{2}\,, \qquad \frac{\partial L_2}{\partial y} = +\frac{\dot{x}}{2} - y\,. \end{aligned} \tag{16.31}$$

The Hamiltonian is now:

$$H(x,y) = \dot{x}p_x + \dot{y}p_y - L_2(x,y;\dot{x},\dot{y}) = \frac{1}{2}\left(\, x^2 + y^2 \,\right), \tag{16.32}$$

and is a function only of x and y. The equations of motion are: $\dot{x} = y$ and $\dot{y} = -x$, from which we can write Hamilton's equations in sympletic form as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix} = \begin{pmatrix} \{\, x, H(x,y)\,\} \\ \{\, x, H(x,y)\,\} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \partial_x H(x,y) \\ \partial_y H(x,y) \end{pmatrix}. \tag{16.33}$$

So we must take x and y as independent sympletic variables. That is from Eq. (16.31), we see that the canonical momentum variables p_x and p_y are *not* independent variables. So we must define Poisson brackets in this case by:

$$\begin{aligned} \{\, A(x,y), B(x,y)\,\} &= \begin{pmatrix} \partial_x A(x,y), & \partial_y A(x,y) \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \partial_x B(x,y) \\ \partial_y B(x,y) \end{pmatrix}, \\ &= \partial_x A(x,y)\,\partial_y B(x,y) - \partial_x B(x,y)\,\partial_y A(x,y)\,. \end{aligned} \tag{16.34}$$

In particular, if $A(x,y) = x$ and $B(x,y) = y$, we have:

$$\{\, x, y\,\} = 1\,. \tag{16.35}$$

Notice that this means that

$$\{\, x, p_x\,\} = \frac{1}{2}\{\, x, y\,\} = \frac{1}{2}\,, \tag{16.36}$$

not one, as would be expected using the canonical momentum p_x as an independent coordinate. So the quantization rule is that $x \mapsto X$ and $y \mapsto Y$, with the Hamiltonian:

$$H(X,Y) = \frac{1}{2}\left(\, X^2 + Y^2 \,\right), \qquad [\, X, Y\,] = i\,, \tag{16.37}$$

which is the same Hamiltonian as before with the same variables and commutation rules. The lesson to be learned here is that one must be careful to identify the independent sympletic variables using Hamilton's classical equations of motion. The Heisenberg equations of motion are:

$$\dot{X} = Y\,, \qquad \dot{Y} = -X\,. \tag{16.38}$$

The Hamiltonian (16.37) is invariant under rotations in the xy-plane, which means that the "angular momentum", defined by:

$$L_z = Y\,\dot{X} - X\,\dot{Y} \tag{16.39}$$

is a constant of the motion. Indeed, from (16.38), we find:

$$L_z = X^2 + Y^2 = 2\,H = 2\,N + 1\,, \tag{16.40}$$

so that $|\, n\,\rangle$ are eigenvectors of L_z with eigenvalues:

$$L_z\,|\, n\,\rangle = (\, 2n + 1\,)\,|\, n\,\rangle\,, \qquad n = 0, 1, 2, \ldots\,. \tag{16.41}$$

That is, we can write: $H = L_z/2$, with the eigenvalues of L_z being $2n + 1 = 1, 3, 5, \ldots$. This suggests a connection between the harmonic oscillator problem and angular momentum theory. In fact, Schwinger exploted this relation between $SU(2)$ and $O(3)$ to compute angular momentum coefficients.

16.4 Coherent states

Another way to view the harmonic oscillator is one that minimizes *both* Δq and Δp. This construction is most closely related to the classical picture of the harmonic oscillator and was originated by Schrödinger in an early article [1]. As we found in Section 1.6.1, the minimization of Heisenberg's uncertainty relation for the harmonic oscillator leads to an eigenvalue equation for the non-Hermitian creation operator A. These are called **coherent states**. Some important papers in the subject are those of Dirac [2], Bargmann [3], Klauder [4], and Klauder and Sudarshan [5].

In our units, the mimimum state is given by:

$$\Delta q = \Delta p = \frac{1}{\sqrt{2}}\,, \qquad \text{with} \qquad \Delta q\,\Delta p = \frac{1}{2}\,. \tag{16.42}$$

From our discussion of the uncertainty principle in Section 1.6.1 on page 25, the minimum $|\,\psi\,\rangle$ is the solution of Eq. (1.132), which in our case becomes:

$$\{\,\Delta p\,Q + i\Delta q\,P\,\}\,|\,\psi\,\rangle = \{\,\Delta p\,\bar{q} + i\,\Delta q\,\bar{p}\,\}\,|\,\psi\,\rangle\,, \tag{16.43}$$

where $\bar{q}$ and $\bar{p}$ are the average values of Q and P. Using (16.42), this becomes:

$$\frac{1}{\sqrt{2}}\,(\,Q + iP\,)\,|\,\psi\,\rangle = \frac{1}{\sqrt{2}}\,(\,\bar{q} + i\bar{p}\,)\,|\,\psi\,\rangle\,. \tag{16.44}$$

But $A = (\,Q + iP\,)/\sqrt{2}$, and if we put $a = (\,\bar{q} + i\bar{p}\,)/\sqrt{2}$, we want to find the ket $|\,a\,\rangle \equiv |\,\psi\,\rangle$ which is the solution of the eigenvalue problem:

$$A\,|\,a\,\rangle = a\,|\,a\,\rangle\,, \tag{16.45}$$

where a is a complex number. Here the operator A is not Hermitian, the eigenvalues a are complex, and the eigenvectors $|\,a\,\rangle$ are not orthogonal. We will sometimes label the vectors $|\,a\,\rangle$ as $|\,\bar{q},\bar{p}\,\rangle$, and sometimes by $|\,a,a^*\,\rangle$, depending on what is needed to describe the state. Let us first find a relation between the eigenvectors $|\,a\,\rangle$ and the eigenvectors of the number operator $|\,n\,\rangle$. We find:

$$\langle\,n\,|\,A\,|\,a\,\rangle = a\,\langle\,n\,|\,a\,\rangle\,. \tag{16.46}$$

Using (16.22), we find:

$$\sqrt{n+1}\,\langle\,n+1\,|\,a\,\rangle = a\,\langle\,n\,|\,a\,\rangle\,, \tag{16.47}$$

from which we find by induction:

$$\langle\,n\,|\,a\,\rangle = \mathcal{N}(a)\,\frac{a^n}{\sqrt{n!}}\,, \qquad \text{where} \qquad \mathcal{N}(a) = \langle\,0\,|\,a\,\rangle\,. \tag{16.48}$$

Using (16.23), this gives:

$$|\,a\,\rangle = \sum_{n=0}^{\infty} |\,n\,\rangle\,\langle\,n\,|\,a\,\rangle = \mathcal{N}(a)\sum_{n=0}^{\infty}\frac{[\,a\,A^\dagger\,]^n}{n!}\,|\,0\,\rangle = \mathcal{N}(a)\,\exp\!\big\{\,a\,A^\dagger\,\big\}\,|\,0\,\rangle\,, \tag{16.49}$$

where $|\,0\,\rangle$ is the $n = 0$ eigenstate of the number operator N. The normalization is arbitrary, but here we choose it such that:

$$\begin{aligned}\langle\,a\,|\,a\,\rangle &= \big|\,\mathcal{N}(a)\,\big|^2\,\langle\,0\,|\,e^{a^*A}\,e^{a\,A^\dagger}\,|\,0\,\rangle = \big|\,\mathcal{N}(a)\,\big|^2\,e^{|\,a\,|^2}\,\langle\,0\,|\,e^{a\,A^\dagger}\,e^{a^*A}\,|\,0\,\rangle \\ &= \big|\,\mathcal{N}(a)\,\big|^2\,e^{|\,a\,|^2} = 1\,, \qquad \text{so} \qquad \mathcal{N}(a) = e^{-|\,a\,|^2/2}\end{aligned} \tag{16.50}$$

where we have used Eq. (B.16) in Appendix B twice. So then

$$|\,a\,\rangle = \exp\!\big\{\,-|\,a\,|^2/2 + a\,A^\dagger\,\big\}\,|\,0\,\rangle = D(a)\,|\,0\,\rangle\,, \tag{16.51}$$

$$D(a) = \exp\!\big\{\,a\,A^\dagger - a^*A\,\big\} = \exp\,i\big\{\,\bar{p}\,Q - \bar{q}\,P\,\big\} = D(\bar{q},\bar{p})\,, \tag{16.52}$$

where $a = (\bar{q} + i\bar{p})/\sqrt{2}$ and $a^* = (\bar{q} - i\bar{p})/\sqrt{2}$. It is easy to show that $D(a)$ is a displacement operator:

$$D^\dagger(a)\, A\, D(a) = A + a\,, \qquad D^\dagger(a)\, A^\dagger\, D(a) = A^\dagger + a^*\,, \tag{16.53}$$
$$D^\dagger(\bar{q},\bar{p})\, Q\, D(\bar{q},\bar{p}) = Q + \bar{q}\,, \qquad D^\dagger(\bar{q},\bar{p})\, P\, D(\bar{q},\bar{p}) = P + \bar{q}\,,$$

The normalization choice of Eq. (16.50) is not the best one to use for some applications. Let us instead take the normalization $\mathcal{N}(a) = 1$, and write the coherent state for this normalization, $|\, a\,)$. Then we have:

$$\langle\, n\,|\, a\,) = \frac{a^n}{\sqrt{n!}}\,, \tag{16.54}$$

so that Eq. (16.49) becomes:

$$|\, a\,) := \sum_{n=0}^{\infty} |\, n\,\rangle\, \langle\, n\,|\, a\,) = \sum_{n=0}^{\infty} \frac{[\, a\, A^\dagger\,]^n}{n!}\, |\, 0\,\rangle = \exp\{\, a\, A^\dagger\,\}\, |\, 0\,\rangle\,, \tag{16.55}$$

so that

$$\left.\frac{\partial^n |\, a\,)}{\partial a^n}\right|_{a=0} = \left[\, A^\dagger\,\right]^n |\, 0\,\rangle = \sqrt{n!}\, |\, n\,\rangle\,. \tag{16.56}$$

So $|\, a\,)$ is a generating function for the harmonic oscillator eigenvectors $|\, n\,\rangle$. For this normalization, we find:

$$A^\dagger\, |\, a\,) = \frac{\partial\, |\, a\,)}{\partial a}\,. \tag{16.57}$$

This exercise shows that a different normalization provides a simple differential representation of $A^\dagger$. Note that $|\, a\,) = e^{|\, a\,|^2/2}\, |\, a\,\rangle$.

Exercise 38. Show that with the normalization choice given in (16.50),

$$\langle\, a'\,|\, a\,\rangle = \exp\{\, -|\, a' - a\,|^2\,\}\,. \tag{16.58}$$

Exercise 39. Show that the average value of the number operator N in a coherent state is:

$$\bar{n} = \langle\, a\,|\, N\,|\, a\,\rangle = |\, a\,|^2\,, \tag{16.59}$$

that the average value of the Hamiltonian for a system in a coherent state is:

$$\langle\, H\,\rangle = \bar{n} + 1/2\,, \tag{16.60}$$

and that for a system in a coherent state, the probability of finding it in an eigenstate of N is given by:

$$P_n(a) = |\, \langle\, n\,|\, a\,\rangle\,|^2 = \frac{\bar{n}^n}{n!}\, e^{-\bar{n}}\,, \tag{16.61}$$

which is a Poisson distribution.

The average values of position and momentum are easily worked out using the same techniques. We find:

$$\begin{aligned}
\bar{q} &= \langle\, Q\,\rangle = \frac{1}{\sqrt{2}}\, \langle\, a\,|\, A + A^\dagger\,|\, a\,\rangle = \frac{1}{\sqrt{2}}\, (a + a^*)\,,\\
\bar{p} &= \langle\, P\,\rangle = \frac{1}{i\sqrt{2}}\, \langle\, a\,|\, A - A^\dagger\,|\, a\,\rangle = \frac{1}{i\sqrt{2}}\, (a - a^*)\,,\\
\overline{q^2} &= \langle\, Q^2\,\rangle = \frac{1}{2}\, \langle\, a\,|\, (A + A^\dagger)^2\,|\, a\,\rangle = \frac{1}{2}\, \langle\, a\,|\, [\, A^\dagger\,]^2 + 2\, A^\dagger A + 1 + [\, A\,]^2\,|\, a\,\rangle\,,\\
&= \frac{1}{2}\, (a + a^*)^2 + \frac{1}{2} = \bar{q}^2 + \frac{1}{2}\,,\\
\overline{p^2} &= \langle\, P^2\,\rangle = -\frac{1}{2}\, \langle\, a\,|\, (A - A^\dagger)^2\,|\, a\,\rangle = -\frac{1}{2}\, \langle\, a\,|\, [\, A^\dagger\,]^2 - 2\, A^\dagger A - 1 + [\, A\,]^2\,|\, a\,\rangle\,,\\
&= -\frac{1}{2}\, (a - a^*)^2 + \frac{1}{2} = \bar{p}^2 + \frac{1}{2}\,,
\end{aligned} \tag{16.62}$$

so that:

$$\Delta q = \sqrt{\overline{q^2} - \bar{q}^2} = 1/\sqrt{2}\,, \qquad \Delta p = \sqrt{\overline{p^2} - \bar{p}^2} = 1/\sqrt{2}\,, \tag{16.63}$$

which is what we assumed to start with.

We can also find a coordinate and momentum representation of a coherent state. First, let us note that by using Eq. (B.16) in Appendix B, we can write $D(\bar{q},\bar{p})$ as:

$$D^\dagger(\bar{q},\bar{p}) = \exp\, i\{\,\bar{q}\,P - \bar{p}\,Q\,\} = \exp\{i\bar{p}\,\bar{q}/2\}\,\exp\{i\bar{q}\,P\}\,\exp\{-i\bar{p}\,Q\}\,. \tag{16.64}$$

Then

$$\begin{aligned}\psi_a(q) &= \langle\, a\,|\,q\,\rangle = \langle\, 0\,|\,D^\dagger(\bar{q},\bar{p})\,|\,q\,\rangle \\ &= \exp\{\,-i(\,\bar{p}\,Q - \bar{p}\,\bar{q}/2\,)\,\}\,\langle\, q\,|\,\exp\{i\bar{q}\,P\}\,|\,0\,\rangle \\ &= \exp\{\,i(\,\bar{p}\,\bar{q} - \bar{p}\,\bar{q}/2\,)\,\}\,\langle\, 0\,|\,q - \bar{q}\,\rangle \\ &= \exp\{\,-[\,q-\bar{q}\,]^2/2 - i\,\bar{p}\,[\,q - \bar{q}/2\,]\,\}/\pi^{1/4}\,.\end{aligned} \tag{16.65}$$

where we have used the normalized ground state solution of Eq. (16.27). So the coherent state is a Gaussian in the coordinate representation of width $1/\sqrt{2}$, centered at $q = \bar{q}$ and with a momentum centered at $p = \bar{p}$, as expected.

In the Heisenberg representation, the displacement operator changes in time according to:

$$\begin{aligned}U^\dagger(t)\,D(a)\,U(t) &= U^\dagger(t)\,\exp\{\,a\,A^\dagger - a^*A\,\}\,U(t) = \exp\{\,a\,A^\dagger(t) - a^*A(t)\,\} \\ &= \exp\{\,a\,A^\dagger\,e^{+it} - a^*A\,e^{-it}\} = \exp\{\,a(t)\,A^\dagger - a^*(t)A\,\} \\ &= D(\,a(t)\,)\,, \qquad \text{where} \qquad a(t) = a\,e^{it}\,.\end{aligned} \tag{16.66}$$

Similarly,

$$U^\dagger(t)\,D(\bar{q},\bar{p})\,U(t) = D(\bar{q}(t),\bar{p}(t))\,, \tag{16.67}$$

where

$$\begin{aligned}\bar{q}(t) &= \bar{q}\,\cos t + \bar{p}\,\sin t\,, \\ \bar{p}(t) &= \bar{p}\,\cos t - \bar{q}\,\sin t\,.\end{aligned} \tag{16.68}$$

So the coherent state eigenvectors change in time according to:

$$\begin{aligned}|\,a,t\,\rangle &= U^\dagger(t)\,|\,a\,\rangle = U^\dagger(t)\,D(a)\,U(t)\,U^\dagger(t)\,|\,0\,\rangle \\ &= e^{it}\,D(\,a(t)\,)|\,0\,\rangle \\ &= e^{it}\,|\,a(t)\,\rangle = e^{it}\,|\,\bar{q}(t),\bar{p}(t)\,\rangle\,,\end{aligned} \tag{16.69}$$

So the time-dependent coordinate representation of a coherent state is given by:

$$\begin{aligned}\psi_a(q,t) &= \langle\, a,t\,|\,q\,\rangle = \langle\, a(t)\,|\,q\,\rangle\,e^{-it} = \langle\, \bar{q}(t),\bar{p}(t)\,|\,q\,\rangle\,e^{-it} \\ &= \exp\{\,-[\,q-\bar{q}(t)\,]^2/2 - i\,\bar{p}(t)\,[\,q - \bar{q}(t)/2\,]\,\}/\pi^{1/4}\,.\end{aligned} \tag{16.70}$$

We have shown here that if the system is in a coherent state, the wave function in coordinate representation is a Gaussian with minimum width centered about the classically oscillating position $q = \bar{q}(t)$, with *no* change in width in either coordinate or momentum representations. That is it moves like a solitary wave, called a **soliton**.

16.4.1 Completeness relations

Coherent states are not orthogonal; however, we can establish a completeness identity for these states. Let us first note that we can write:

$$a = \frac{1}{\sqrt{2}}\left(q/b + ib\,p/\hbar\right), \qquad a^* = \frac{1}{\sqrt{2}}\left(q/b - ib\,p/\hbar\right), \tag{16.71}$$

where q and p are in ordinary units. We can also put $a = \rho e^{i\phi}$. So we find:

$$\frac{\mathrm{d}a\,\mathrm{d}a^*}{2\pi i} = \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar} = \frac{\rho\,\mathrm{d}\rho\,\mathrm{d}\phi}{\pi}\,. \tag{16.72}$$

So we find the completeness relation:

$$\begin{aligned}\iint_{-\infty}^{+\infty}\frac{da\,da^*}{2\pi i}\,|\,a\,\rangle\langle\,a\,| &= \int_0^\infty\int_0^{2\pi}\frac{\rho\,d\rho\,d\phi}{\pi}\,|\,a\,\rangle\langle\,a\,|\,,\\ &= \int_0^\infty\int_0^{2\pi}\frac{\rho\,d\rho\,d\phi}{\pi}\sum_{n,n'=0}^{\infty}|\,n\,\rangle\,\frac{\rho^{n+n'}e^{-\rho^2}e^{i(n-n')\phi}}{\sqrt{n!\,n'!}}\,\langle\,n'\,|\,,\\ &= \sum_{n=0}^{\infty}\frac{|\,n\,\rangle\langle\,n\,|}{n!}\,2\int_0^\infty e^{-\rho^2}\rho^{2n+1}d\rho = \sum_{n=0}^{\infty}|\,n\,\rangle\langle\,n\,| = 1\,.\end{aligned} \tag{16.73}$$

where the integration goes over the entire complex plane. For example a vector $|\,\psi\,\rangle$ can be expanded in coherent states $|\,a\,\rangle$ using (16.73). We find:

$$|\,\psi\,\rangle = \iint_{-\infty}^{+\infty}\frac{da\,da^*}{2\pi i}\,|\,a\,\rangle\,\psi(a)\,, \qquad \psi(a) = \langle\,a\,|\,\psi\,\rangle\,. \tag{16.74}$$

We can also find a trace using coherent states. It is easy to show that:

$$\mathrm{Tr}[\,M\,] = \sum_{n=0}^{\infty}\langle\,n\,|\,M\,|\,n\,\rangle = \iint_{-\infty}^{+\infty}\frac{da\,da^*}{2\pi i}\,\langle\,a\,|\,M\,|\,a\,\rangle = \iint_{-\infty}^{+\infty}\frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,\langle\,q,p\,|\,M\,|\,q,p\,\rangle\,. \tag{16.75}$$

where M is any operator.

16.4.2 Generating function

One of the uses of coherent vectors are as generating functions for matrix elements of operators. As an example, we will compute matrix elements of the operator $[\,A^\dagger\,]^k$. So let us consider:

$$e^{\lambda A^\dagger}\,|\,a,a^*\,\rangle = e^{\lambda A^\dagger}\,e^{a\,A^\dagger - a^*\,A}\,|\,0\,\rangle = e^{\lambda\,a^*/2}\,e^{(a+\lambda)\,A^\dagger - a^*\,A}\,|\,0\,\rangle = e^{\lambda\,a^*/2}\,|\,a+\lambda,a^*\,\rangle\,. \tag{16.76}$$

Operating on the left by $\langle\,n\,|$, and inserting a complete set of states, we find:

$$\sum_{n'=0}^{\infty}\langle\,n\,|\,e^{\lambda A^\dagger}\,|\,n'\,\rangle\,\langle\,n'\,|\,a,a^*\,\rangle = e^{\lambda\,a^*/2}\,\langle\,n\,|\,a+\lambda,a^*\,\rangle\,. \tag{16.77}$$

From Eq. (16.48), we have:

$$\langle\,n\,|\,a,a^*\,\rangle = e^{-a\,a^*/2}\,\frac{a^n}{\sqrt{n!}}\,, \tag{16.78}$$

so (16.77) becomes:

$$e^{-a\,a^*/2}\sum_{n'=0}^{\infty}\langle\,n\,|\,e^{\lambda A^\dagger}\,|\,n'\,\rangle\,\frac{a^{n'}}{\sqrt{n'!}} = e^{[\,-(a+\lambda)\,a^*+\lambda\,a^*\,]/2}\,\frac{(\,a+\lambda\,)^n}{\sqrt{n!}}\,. \tag{16.79}$$

The exponential normalization factors cancel here, as they must. Expanding the left and right sides of this equation in powers of λ using the binomial theorem gives:

$$\sum_{n'=0}^{\infty}\sum_{k=0}^{\infty}\langle\, n\,|\,[\,A^\dagger\,]^k\,|\,n'\,\rangle\,\frac{\lambda^k\,a^{n'}}{k!\,\sqrt{n'!}} = \sum_{n'=0}^{n}\frac{\sqrt{n!}\;a^{n'}\,\lambda^{n-n'}}{(n-n')!\,n'!}\;, \tag{16.80}$$

and comparing coefficients of powers of λ give:

$$\langle\, n\,|\,[\,A^\dagger\,]^k\,|\,n'\,\rangle = \delta_{k,n-n'}\,\sqrt{\frac{n!}{n'!}}\;. \tag{16.81}$$

So coherent states can be used as generating functions for matrix elements of operators.

16.5 Squeezed states

Squeezed states are coherent states with arbitrary values of either Δq *or* Δp, but with minimum value of the product of the two. They can be generated from the coherent states we found in the last section by a unitary transformation to new operators B and $B^\dagger$. We put:

$$\begin{aligned} A &= \lambda\,B + \nu\,B^\dagger\,,\\ A^\dagger &= \lambda^*\,A^\dagger + \nu^*\,A\,, \end{aligned} \tag{16.82}$$

and require that the commutation relations are preserved:

$$[\,A,A^\dagger\,] = (\,|\,\lambda\,|^2 - |\,\nu\,|^2\,)\,[\,B,B^\dagger\,] = |\,\lambda\,|^2 - |\,\nu\,|^2 = 1\,. \tag{16.83}$$

This change of basis is called a **Bogoliubov** transformation, after the fellow who first discovered it. We put:[2]

$$\lambda = \cosh(r)\,,\qquad \nu = e^{i\phi}\,\sinh(r)\,,$$

with r and ϕ real. Then Eq. (16.82) can be written as:

$$\begin{aligned} A &= \cosh(r)\,B + e^{+i\phi}\,\sinh(r)\,B^\dagger = V^\dagger(r,\phi)\,B\,V(r,\phi)\,,\\ A^\dagger &= \cosh(r)\,B^\dagger + e^{-i\phi}\,\sinh(r)\,B = V^\dagger(r,\phi)\,B^\dagger\,V(r,\phi)\,. \end{aligned} \tag{16.84}$$

Using the identities in Appendix B, we easily find:

$$V(z) = V(r,\phi) = \exp\{\,(z\,A^{\dagger\,2} - z^*\,A^2\,)/2\,\} = \exp\{\,(z\,B^{\dagger\,2} - z^*\,B^2\,)/2\,\}\,, \tag{16.85}$$

where we have put:

$$z = r\,e^{i\phi}\,. \tag{16.86}$$

The operator $V(z)$ is called the **squeeze** operator, with squeeze parameter z.

Exercise 40. Show that $V(z)\,V^\dagger(z) = 1$, and that:

$$z^*\,A^2 - z\,A^{\dagger\,2} = z^*\,B^2 - z\,B^{\dagger\,2}\,, \tag{16.87}$$

Exercise 41. Show that the inverse relation of (16.84) is given by:

$$\begin{aligned} B &= \cosh(r)\,A - e^{+i\phi}\,\sinh(r)\,A^\dagger = V(r,\phi)\,A\,V^\dagger(r,\phi)\,,\\ B^\dagger &= \cosh(r)\,A^\dagger - e^{-i\phi}\,\sinh(r)\,A = V(r,\phi)\,A^\dagger\,V^\dagger(r,\phi)\,. \end{aligned} \tag{16.88}$$

[2]The *overall* phase of the transformation is not physically significant.

The Hamiltonian is now given by:

$$H = \frac{1}{2}\left\{ A^\dagger A + AA^\dagger \right\} = \frac{1}{2}\left\{ B^\dagger B + BB^\dagger + \sinh(2r)\left(e^{+i\phi}\, B^{\dagger\,2} + e^{-i\phi}\, B^2 \right) \right\}, \tag{16.89}$$

So Heisenberg's equations of motion for $A(t)$ gives:

$$\begin{aligned} A(t) = U^\dagger(t)\, A\, U(t) = A\, e^{-it} &= \left\{ \cosh(r)\, B + e^{+i\phi}\, \sinh(r)\, B^\dagger \right\} e^{-it}\,, \\ A^\dagger(t) = U^\dagger(t)\, A^\dagger\, U(t) = A^\dagger\, e^{+it} &= \left\{ \cosh(r)\, B^\dagger + e^{-i\phi}\, \sinh(r)\, B \right\} e^{+it}\,, \end{aligned} \tag{16.90}$$

whereas for $B(t)$, using Eq. (16.88), we find:

$$\begin{aligned} B(t) = U^\dagger(t)\, B\, U(t) &= \cosh(r)\, A(t) - e^{+i\phi}\, \sinh(r)\, A^\dagger(t) \\ &= \cosh(r)\, A\, e^{-it} - e^{+i\phi}\, \sinh(r)\, A^\dagger\, e^{+it}\,, \\ B^\dagger(t) = U^\dagger(t)\, B^\dagger\, U(t) &= \cosh(r)\, A^\dagger(t) - e^{-i\phi}\, \sinh(r)\, A(t) \\ &= \cosh(r)\, A^\dagger\, e^{+it} - e^{-i\phi}\, \sinh(r)\, A\, e^{-it}\,. \end{aligned} \tag{16.91}$$

Now let $|\,a\,\rangle_b$ be a coherent eigenvector of B with complex eigenvalue a satisfying:

$$B\,|\,a\,\rangle_b = a\,|\,a\,\rangle_b\,. \tag{16.92}$$

Multiplying on the left by $V^\dagger(z)$ and using (16.84), we find:

$$V^\dagger(z)\, B\, V(z)\, V^\dagger(z)\,|\,a\,\rangle_b = A\, V^\dagger(z)\,|\,a\,\rangle_b = a\, V^\dagger(z)\,|\,a\,\rangle_b\,, \tag{16.93}$$

so that $V^\dagger(z)\,|\,a\,\rangle_b$ is an eigenvector A with eigenvalue a. Solving for $|\,a\,\rangle_b$, we find:

$$|\,a\,\rangle_b = V(z)\,|\,a\,\rangle_a\,. \tag{16.94}$$

For the system in a squeezed state $|\,a\,\rangle_b$, the average values of Q and P are given by:

$$\begin{aligned} \bar{q}_b(t) = {}_b\langle\, a\,|\, Q(t)\,|\,a\,\rangle_b &= \frac{1}{i\sqrt{2}}\, {}_b\langle\, a\,|\,(\, A(t) + A^\dagger(t)\,)|\,a\,\rangle_b \\ = \frac{1}{i\sqrt{2}}\, {}_b\langle\, a\,|\, \Big\{ \left\{ \cosh(r)\, B + e^{+i\phi}\, \sinh(r)\, B^\dagger \right\} e^{-it} &+ \left\{ \cosh(r)\, B^\dagger + e^{-i\phi}\, \sinh(r)\, B \right\} e^{+it} \Big\}\,|\,a\,\rangle_b \\ = \cosh(r)\left(\bar{q}_a\, \cos(t) + \bar{p}_a\, \sin(t) \right) &+ \sinh(r)\left(\bar{q}_a\, \cos(t-\phi) - \bar{p}_a\, \sin(t-\phi) \right), \\ = q_0\, \cos(t) + p_0\, \sin(t)\,, & \end{aligned} \tag{16.95}$$

and

$$\begin{aligned} \bar{p}_b(t) = {}_b\langle\, a\,|\, P(t)\,|\,a\,\rangle_b &= \frac{1}{i\sqrt{2}}\, {}_b\langle\, a\,|\,(\, A(t) - A^\dagger(t)\,)|\,a\,\rangle_b \\ = \frac{1}{\sqrt{2}}\, {}_b\langle\, a\,|\, \Big\{ \left\{ \cosh(r)\, B + e^{+i\phi}\, \sinh(r)\, B^\dagger \right\} e^{-it} &- \left\{ \cosh(r)\, B^\dagger + e^{-i\phi}\, \sinh(r)\, B \right\} e^{+it} \Big\}\,|\,a\,\rangle_b \\ = \cosh(r)\left(\bar{p}_a\, \cos(t) - \bar{q}_a\, \sin(t) \right) &- \sinh(r)\left(\bar{p}_a\, \cos(t-\phi) + \bar{q}_a\, \sin(t-\phi) \right) \\ = p_0\, \cos(t) - q_0\, \sin t\,, & \end{aligned} \tag{16.96}$$

where

$$\begin{aligned} q_0 &= \bar{q}_a\, \cosh(r) + \sinh(r)\,(\, \bar{q}_a\, \cos\phi + \bar{p}_a\, \sin\phi\,)\,, \\ p_0 &= \bar{p}_a\, \cosh(r)\, + \sinh(r)\,(\, \bar{p}_a\, \cos\phi - \bar{q}_a\, \sin\phi\,)\,, \\ \bar{q}_a &= (a + a^*)/\sqrt{2}\,, \qquad \bar{p}_a = (a - a^*)/i\sqrt{2}\,. \end{aligned}$$

Note that $\bar{q}_b(t)$ and $\bar{p}_b(t)$ oscillate with the classical frequency, and that $\dot{\bar{q}}_b(t) = \bar{p}_b(t)$ and $\dot{\bar{p}}_b(t) = -\bar{q}_b(t)$, as required for the classical solution. For the width functions, we find:

$$\begin{aligned}
\overline{q^2}_b(t) &= {}_b\langle\, a \,|\, Q^2(t) \,|\, a \,\rangle_b = \frac{1}{2}\, {}_b\langle\, a \,|(\, A(t) + A^\dagger(t)\,)^2|\, a \,\rangle_b \\
&= \bar{q}_b^2(t) + \frac{1}{2}\left\{\, \cosh(2r) + \sinh(2r)\, \cos(2t-\phi)\,\right\}, \\
\overline{p^2}_b(t) &= {}_b\langle\, a \,|\, P^2(t) \,|\, a \,\rangle_b = -\frac{1}{2}\, {}_b\langle\, a \,|(\, A(t) - A^\dagger(t)\,)^2|\, a \,\rangle_b \\
&= \bar{p}_b^2(t) + \frac{1}{2}\left\{\, \cosh(2r) - \sinh(2r)\, \cos(2t-\phi)\,\right\},
\end{aligned} \tag{16.97}$$

so the width functions are given by:

$$\begin{aligned}
[\,\Delta q_b(t)\,]^2 &= \left\{\, \cosh(2r) + \sinh(2r)\, \cos(2t-\phi)\,\right\}/2\,, \\
[\,\Delta p_b(t)\,]^2 &= \left\{\, \cosh(2r) - \sinh(2r)\, \cos(2t-\phi)\,\right\}/2\,.
\end{aligned} \tag{16.98}$$

The uncertainty product is:

$$[\,\Delta q_b(t)\,]^2\,[\,\Delta p_b(t)\,]^2 = \left\{\, 1 + \sinh^2(2r)\, \sin^2(2t-\phi)\,\right\}/4\,. \tag{16.99}$$

So if r is very large, Δq_b and Δp_b oscillate between very large values and very small values with a frequency of twice the natural frequency of the oscillator.

The time dependent coordinate representation for the squeezed state is more difficult to find. The time dependence of the squeeze operator is given by:

$$\begin{aligned}
U^\dagger(t)\, V(z)\, U(t) &= \exp\!\left\{\, (z\, A^{\dagger\,2}(t) - z^*\, A^2(t)\,)/2\,\right\} \\
&= \exp\!\left\{\, (z\, A^{\dagger\,2}\, e^{+2it} - z^*\, A^2\, e^{-2it}\,)/2\,\right\} \\
&= \exp\!\left\{\, (z(t)\, A^{\dagger\,2} - z^*(t)\, A^2\,)/2\,\right\} \\
&= \exp\!\left\{\, (z(t)\, B^{\dagger\,2} - z^*(t)\, B^2\,)/2\,\right\} \\
&= V(\, z(t)\,)\,,
\end{aligned} \tag{16.100}$$

where

$$z(t) = r\, e^{i\phi(t)}\,, \qquad \phi(t) = \phi + 2t\,. \tag{16.101}$$

Thus only $\phi(t)$ depends linearly on t. The Heisenberg time dependent squeezed state is given by:

$$|\, a,t\,\rangle_b = U^\dagger(t)\,|\, a\,\rangle_b = U^\dagger(t)\, V(z)\, U(t)\, U^\dagger(t)\,|\, a\,\rangle_a = V(z(t))\,|\, a,t\,\rangle_a\,. \tag{16.102}$$

But since $A = (Q + iP)/\sqrt{2}$ and $A^\dagger = (Q - iP)/\sqrt{2}$, we have:

$$\begin{aligned}
V^\dagger(z(t)) &= \exp\!\left\{\, (\, z^*(t)\, A^2 - z(t)\, A^{\dagger\,2}\,)/2\,\right\} \\
&= \exp\!\left\{\, -ir\left\{\, \sin(\phi(t))\,(\, Q^2 + P^2\,) + \cos(\phi(t))\,(\, QP + PQ\,)\,\right\}\right\}.
\end{aligned} \tag{16.103}$$

So we will need to find:

$$\psi_a(q,t) = \langle\, a,t\,|\, q\,\rangle = \langle\, a,t\,|\, V^\dagger(z(t))\,|\, q\,\rangle = \langle\, 0\,|\, D(a(t))\, V^\dagger(z(t))\,|\, q\,\rangle\,. \tag{16.104}$$

Squeezed states have been extensively studied in the literature [6]. In fact, Nieto points out in his review that the squeezed state wave function was first found by Schrödinger in 1926 [1]in his attempts to construct a wave theory of quantum mechanics, and by Kennard in 1927 [7]. These exact solutions of Schrödinger's equation for the harmonic oscillator were interesting at the time because they tracked the classical motion as closely as possible. The interest in squeezed states was revived in the 1980's when optical squeezed-state lasers were constructed.

16.6 The forced oscillator

The Hamiltonian for a harmonic oscillator driven by an external force $f(t)$ is given by:

$$H(Q,P) = \frac{P^2}{2m} + \frac{1}{2}m\,\omega_0^2\,Q^2 - Qf(t)\,, \qquad \text{where} \qquad [\,Q,P\,] = i\hbar\,, \tag{16.105}$$

and where $f(t)$ commutes with all quantum operators. Heisenberg's equations of motion for this system are given by:

$$\begin{aligned} \dot{Q} &= [\,Q,H\,]/i\hbar = P/m\,, \\ \dot{P} &= [\,P,H\,]/i\hbar = -m\,\omega_0^2\,Q + f(t)\,, \end{aligned} \tag{16.106}$$

from which we find:

$$\ddot{Q} + \omega_0^2\,Q = j(t)\,, \tag{16.107}$$

where $j(t) = f(t)/m$. Let us first study a problem where $|j(t)| \to 0$ as $t \to \pm\infty$. Then in both these limits, $Q(t)$ satisfies a homogenious equation. We call the solutions for $t \to -\infty$ the "in" operators and the solutions for $t \to +\infty$ the "out" operators. They both satisfy a homogenious equation:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega_0^2\right\}\begin{pmatrix} Q_{\rm in}\,(t) \\ Q_{\rm out}(t)\end{pmatrix} = 0\,, \tag{16.108}$$

with equal time commutation relations:

$$[\,Q_{\rm in}\,(t), P_{\rm in}\,(t)\,] = i\hbar\,, \qquad [\,Q_{\rm out}(t), P_{\rm out}(t)\,] = i\hbar\,. \tag{16.109}$$

Solutions of (16.108) are given by:

$$\begin{pmatrix} Q_{\rm in}\,(t) \\ Q_{\rm out}(t)\end{pmatrix} = \sqrt{\frac{\hbar}{2m\omega_0}}\left\{\begin{pmatrix} A_{\rm in} \\ A_{\rm out}\end{pmatrix} e^{-i\omega_0 t} + \begin{pmatrix} A^\dagger_{\rm in} \\ A^\dagger_{\rm out}\end{pmatrix} e^{i\omega_0 t}\right\}\,. \tag{16.110}$$

The momentum operators are given by:

$$\begin{pmatrix} P_{\rm in}\,(t) \\ P_{\rm out}(t)\end{pmatrix} = \frac{1}{i}\sqrt{\frac{\hbar m\omega_0}{2}}\left\{\begin{pmatrix} A_{\rm in} \\ A_{\rm out}\end{pmatrix} e^{-i\omega_0 t} - \begin{pmatrix} A^\dagger_{\rm in} \\ A^\dagger_{\rm out}\end{pmatrix} e^{i\omega_0 t}\right\}\,. \tag{16.111}$$

With our choice of normalization constants, the commutation relations (16.109) require that:

$$[\,A_{\rm in}\,, A^\dagger_{\rm in}\,] = 1\,, \qquad [\,A_{\rm out}, A^\dagger_{\rm out}\,] = 1\,. \tag{16.112}$$

The operators $A_{\rm in}$ and $A_{\rm out}$ are *different* operators in Hilbert space but they have the same commutation relations. The in and out Hamiltonians have the same form:

$$\begin{aligned} H_{\rm in} \;&= H(-\infty) = \hbar\omega_0\left\{A^\dagger_{\rm in}\,A_{\rm in}\; + \frac{1}{2}\right\}\,, \\ H_{\rm out} &= H(+\infty) = \hbar\omega_0\left\{A^\dagger_{\rm out}A_{\rm out} + \frac{1}{2}\right\}\,, \end{aligned} \tag{16.113}$$

and have the same eigenvalue spectrum. The eigenvalue equations for the in and out Hamiltonians are written as:

$$\begin{aligned} H_{\rm in}\,|\,n\,\rangle_{\rm in} \;&= E_n|\,n\,\rangle_{\rm in}\;\,, \\ H_{\rm out}|\,n\,\rangle_{\rm out} &= E_n|\,n\,\rangle_{\rm out}\,, \end{aligned} \tag{16.114}$$

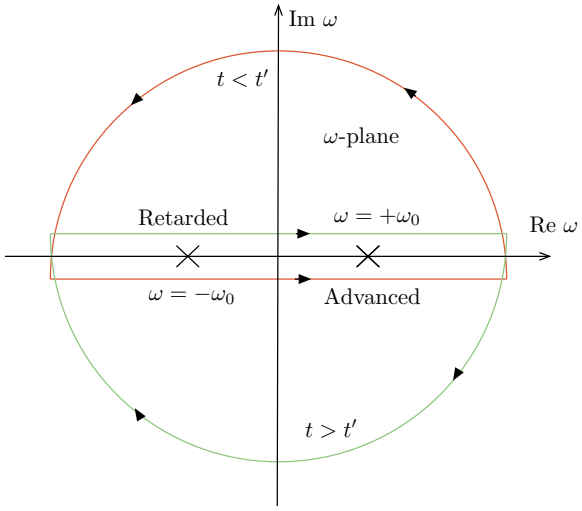

Figure 16.1: Retarded and advanced contours for the Green function of Eq. (16.117).

Where $E_n = \hbar\omega(n + 1/2)$. Both of these states are complete sets of states for the physical system, and so they must be related by a unitary transformation, which we call S. That is, we can write:

$$A_{\rm out} = S^\dagger\, A_{\rm in}\, S\,, \qquad \text{and} \qquad A^\dagger_{\rm out} = S^\dagger\, A^\dagger_{\rm in}\, S\,, \tag{16.115}$$

and

$$|\,n\,\rangle_{\rm out} = S^\dagger\,|\,n\,\rangle_{\rm in}\,, \qquad \text{and} \qquad {}_{\rm out}\langle\, n\,|\,n\,\rangle_{\rm in} = {}_{\rm out}\langle\, n\,|\,S\,|\,n\,\rangle_{\rm in}\,.$$

If all we care about is the relation between in- and out-states, our problem is to find S. We do this by finding solutions to Eq. (16.107) which reduce to in and out states when $t \to \mp\infty$. This means we need to find retarded and advanced solutions to the Green function equation:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega_0^2\right\} G(t,t') = \delta(t-t')\,. \tag{16.116}$$

We put:

$$G(t,t') = \int_{\mathcal{C}} \frac{\mathrm{d}\omega}{2\pi}\,\tilde{G}(\omega)\,e^{-i\omega(t-t')}\,, \tag{16.117}$$

where $\mathcal{C}$ is a contour, to be specified, which runs from $\omega = -\infty$ to $\omega = +\infty$ along the real axis. Then Eq. (16.116) is satisfied if

$$\tilde{G}(\omega) = \frac{1}{\omega_0^2 - \omega^2}\,. \tag{16.118}$$

So $\tilde{G}(\omega)$ is analytic everywhere, except for simple poles at $\omega = \pm\omega_0$. Contours for retarded and advanced solutions are shown in Fig. 16.1, and we find:

$$\begin{aligned} G_R(t,t') &= \frac{i}{2\omega_0}\left\{e^{-i\omega_0(t-t')} - e^{i\omega_0(t-t')}\right\}\Theta(t-t')\,,\\ G_A(t,t') &= \frac{-i}{2\omega_0}\left\{e^{-i\omega_0(t-t')} - e^{i\omega_0(t-t')}\right\}\Theta(t'-t)\,. \end{aligned} \tag{16.119}$$

We also find that:

$$G_R(t,t') - G_A(t,t') = \frac{i}{2\omega_0}\left\{e^{-i\omega_0(t-t')} - e^{i\omega_0(t-t')}\right\}\,. \tag{16.120}$$

So solutions which reduce to in and out states at $t \to \mp\infty$ are given by:

$$Q(t) = Q_{\rm in}\,(t) + \int_{-\infty}^{+\infty} G_R(t,t')\,j(t')\,{\rm d}t' = Q_{\rm out}(t) + \int_{-\infty}^{+\infty} G_A(t,t')\,j(t')\,{\rm d}t'\,, \tag{16.121}$$

from which we find:

$$Q_{\rm out}(t) = Q_{\rm in}\,(t) + \int_{-\infty}^{+\infty} \{\, G_R(t,t') - G_A(t,t')\,\}\,j(t')\,{\rm d}t'\,, \tag{16.122}$$

$$= Q_{\rm in}\,(t) + \frac{i}{2\omega_0}\{\,\tilde{j}(\omega_0)\,e^{-i\omega_0 t} - \tilde{j}^*(\omega_0)\,e^{i\omega_0 t}\,\}\,, \tag{16.123}$$

where $\tilde{j}(\omega_0)$ is the Fourier transform of the current, evaluated at the oscillator resonant frequency,

$$\tilde{j}(\omega_0) = \int_{-\infty}^{\infty} j(t)\,e^{i\omega_0 t}\,{\rm d}t\,. \tag{16.124}$$

Using (16.110), we find:

$$\begin{aligned} A_{\rm out} &= A_{\rm in}\ + a = S^\dagger\,A_{\rm in}\,S\,,\\ A^\dagger_{\rm out} &= A^\dagger_{\rm in}\ + a^* = S^\dagger\,A^\dagger_{\rm in}\,S\,, \end{aligned} \tag{16.125}$$

where a is a c-number, given by:

$$a = i\sqrt{\frac{m}{2\hbar\omega_0}}\ \tilde{j}(\omega_0)\,. \tag{16.126}$$

Note that a depends only on the Fourier transform of the driving force evaluated at the resonate frequency: $\omega = \omega_0$. The unitary operator S, which generates (16.125) is given by a displacement operator:

$$S(a) = \exp\{\,a\,A^\dagger_{\rm in} - a^*\,A_{\rm in}\,\} \equiv D(a)\,, \tag{16.127}$$

where $D(a)$ is the operator given in Eq. (16.52) discussed in Section 16.4 on coherent states.

Now suppose that the system is in the ground state $|\,0\,\rangle_{\rm in}$ of $H_{\rm in}$ at $t \to -\infty$. Then the final state is given by the coherent state:

$$|\,a\,\rangle_{\rm out} = S^\dagger(a)\,|\,0\,\rangle_{\rm in} = D(-a)\,|\,0\,\rangle_{\rm in}\,, \tag{16.128}$$

and the probability of finding the system in the state $|\,n\,\rangle$ is given by the coherent state Poisson probability amplitude:

$$P_n(a) = |\,{}_{\rm out}\langle\,n\,|\,0\,\rangle_{\rm in}\,|^2 = \frac{\bar{n}^n\,e^{-\bar{n}}}{n!}\,, \qquad \bar{n} = |\,a\,|^2 = \frac{m}{2\hbar\omega_0}\ \left|\,\tilde{j}(\omega_0)\,\right|^2\,. \tag{16.129}$$

That is, if we start with the sytem in the ground state at $t \to -\infty$, then at $t \to +\infty$, the system will be in a coherent state with the average value of $\langle\,N\,\rangle = \bar{n}$, distributed with a Poisson distribution about this average value. The system started out in a pure state of the oscillator, the ground state, with energy $E_{\rm in} = \hbar\omega_0/2$. The average energy in the final state at $t \to +\infty$ is $E_{\rm out} = \hbar\omega_0\,(\,\bar{n} + 1/2\,)$. Thus energy has been pumped into the system by the applied force: energy is not conserved. On the average, the total work done on the system by the applied force is given by:

$$W = E_{\rm out} - E_{\rm in} = \hbar\omega_0\,\bar{n} = \frac{m}{2}\,|\tilde{j}(\omega_0)|^2 = \frac{|\tilde{f}(\omega_0)|^2}{2m}\,, \tag{16.130}$$

where $\tilde{f}(\omega_0)$ is the Fourier transform of the external force. The work done W is independent of $\hbar$, and so must agree with the *classical* result for the energy transfer. (See Exercise 44 below.)

We can also use Green functions which have boundary conditions at both $t = +\infty$ and $t = -\infty$. These are called the Feynman and anti-Feynman Green functions and are defined by the contours shown in Fig. 16.2.

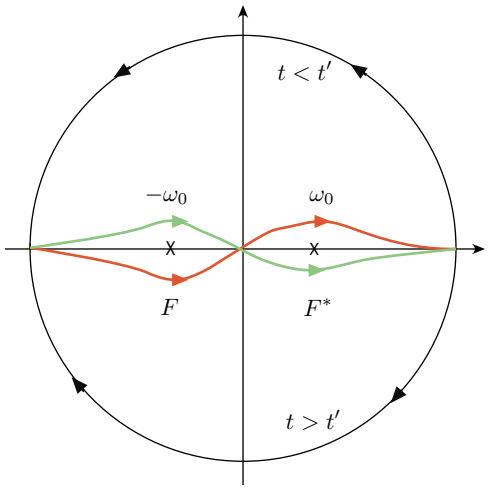

Figure 16.2: Feynman (F) (red) and anti-Feynman (F^*) (green) contours.

Exercise 42. Show that:

$$\begin{aligned} G_F(t-t') &= -\int_F \frac{\mathrm{d}\omega}{2\pi}\,\frac{e^{-i\omega(t-t')}}{\omega^2-\omega_0^2} = \frac{i}{2\omega_0}\left\{ e^{-i\omega_0(t-t')}\,\theta(t-t') + e^{i\omega_0(t-t')}\,\theta(t'-t)\right\}, \\ G_{F^*}(t-t') &= -\int_{F^*} \frac{\mathrm{d}\omega}{2\pi}\,\frac{e^{-i\omega(t-t')}}{\omega^2-\omega_0^2} = \frac{-i}{2\omega_0}\left\{ e^{i\omega_0(t-t')}\,\theta(t-t') + e^{-i\omega_0(t-t')}\,\theta(t'-t)\right\}, \end{aligned} \tag{16.131}$$

and that

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2}+\omega_0^2\right\} G_{F,F^*}(t,t') = \delta(t-t')\,. \tag{16.132}$$

Exercise 43. Show that for the Feynman and anti-Feynman Green functions, the solution for $Q(t)$ and $Q^\dagger(t)$ can be written as:

$$\begin{aligned} Q(t) &= Q_0(t) + \int_{-\infty}^{+\infty} G_F(t-t')\,j(t')\,\mathrm{d}t'\,, \\ \text{and} \qquad Q^\dagger(t) &= Q_0^\dagger(t) + \int_{-\infty}^{+\infty} G_{F^*}(t-t')\,j(t')\,\mathrm{d}t'\,, \end{aligned} \tag{16.133}$$

where

$$Q_0(t) = \sqrt{\frac{\hbar}{2\omega_0 m}}\left\{ A_{\mathrm{in}}\,e^{-i\omega_0 t} + A_{\mathrm{out}}^\dagger\,e^{-i\omega_0 t}\right\}. \tag{16.134}$$

(Note $j(t)$ is *real*.) Using the fact that

$$\lim_{t\to+\infty} Q(t) = Q_{\mathrm{out}}(t)\,, \qquad \text{and} \qquad \lim_{t\to-\infty} Q(t) = Q_{\mathrm{in}}(t)\,, \tag{16.135}$$

show that using these Green functions, we again find that:

$$A_{\mathrm{out}} = A_{\mathrm{in}} + a\,, \qquad \text{and} \qquad A_{\mathrm{out}}^\dagger = A_{\mathrm{in}}^\dagger + a^*\,, \tag{16.136}$$

where a is given by Eq. (16.126), so that (16.136) is in agreement with Eq. (16.125).

Exercise 44. The work done by a force $f(t)$ on a classical harmonic oscillator is given by:

$$W = \int_{-\infty}^{+\infty} f(t)\,\dot{q}(t)\,\mathrm{d}t = m\int_{-\infty}^{+\infty} j(t)\,\dot{q}(t)\,\mathrm{d}t\,, \tag{16.137}$$

where $j(t) = f(t)/m$. Calculate the total work done if the oscillator starts at rest at $t \to -\infty$, and show that it agrees with Eq. (16.130). [Hint: use the retarded Green function.]

Solution: The equation of motion for the driven oscillator is:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega_0^2\right\} q(t) = j(t)\,, \tag{16.138}$$

with $q(t) \to 0$ and $\dot{q}(t) \to 0$ as $t \to -\infty$. So it will be useful to use a retarded Green function here, and write the solution as:

$$q(t) = q_0(t) + \int_{-\infty}^{+\infty} G_R(t-t')\,j(t')\,\mathrm{d}t'\,, \tag{16.139}$$

with $q_0(t) = 0$ and where the retarded Green function is given by:

$$G_R(t-t') = \frac{i}{2\omega_0}\left\{e^{-i\omega_0(t-t')} - e^{+i\omega_0(t-t')}\right\}\Theta(t-t')\,. \tag{16.140}$$

So

$$\begin{aligned}\dot{q}(t) &= \int_{-\infty}^{+\infty}\frac{\mathrm{d}G_R(t-t')}{\mathrm{d}t}\,j(t')\,\mathrm{d}t' \\ &= \frac{1}{2}\int_{-\infty}^{t}\left\{e^{-i\omega_0(t-t')} + e^{+i\omega_0(t-t')}\right\} j(t')\,\mathrm{d}t'\,.\end{aligned} \tag{16.141}$$

Substitution into (16.137) gives:

$$\begin{aligned}W &= m\int_{-\infty}^{+\infty} j(t)\,\dot{q}(t)\,\mathrm{d}t \\ &= \frac{m}{2}\int_{-\infty}^{+\infty}\mathrm{d}t\,j(t)\int_{-\infty}^{t}\mathrm{d}t'\,\left\{e^{-i\omega_0(t-t')} + e^{+i\omega_0(t-t')}\right\} j(t') \\ &= \frac{m}{2}\left\{\int_{-\infty}^{+\infty}\mathrm{d}t\,j(t)\,e^{-i\omega_0 t}\int_{-\infty}^{t}\mathrm{d}t'\,j(t')\,e^{+i\omega_0 t'} + \int_{-\infty}^{+\infty}\mathrm{d}t\,j(t)\,e^{+i\omega_0 t}\int_{-\infty}^{t}\mathrm{d}t'\,j(t')\,e^{-i\omega_0 t'}\right\}.\end{aligned} \tag{16.142}$$

In the second term, first interchange t and t'. Then change the order of integration, keeping in mind the region of integration. This gives:

$$\begin{aligned}\int_{-\infty}^{+\infty}\mathrm{d}t\,j(t)\,e^{+i\omega_0 t}\int_{-\infty}^{t}\mathrm{d}t'\,j(t')\,e^{-i\omega_0 t'} &= \int_{-\infty}^{+\infty}\mathrm{d}t'\,j(t')\,e^{+i\omega_0 t'}\int_{-\infty}^{t'}\mathrm{d}t\,j(t)\,e^{-i\omega_0 t} \\ &= \int_{-\infty}^{+\infty}\mathrm{d}t\,j(t)\,e^{-i\omega_0 t}\int_{t}^{+\infty}\mathrm{d}t'\,j(t')\,e^{+i\omega_0 t'}\,.\end{aligned} \tag{16.143}$$

Substituting this into (16.142) gives:

$$\begin{aligned}W &= \frac{m}{2}\left\{\int_{-\infty}^{+\infty}\mathrm{d}t\,j(t)\,e^{-i\omega_0 t}\int_{-\infty}^{t}\mathrm{d}t'\,j(t')\,e^{+i\omega_0 t'} + \int_{-\infty}^{+\infty}\mathrm{d}t\,j(t)\,e^{-i\omega_0 t}\int_{t}^{+\infty}\mathrm{d}t'\,j(t')\,e^{+i\omega_0 t'}\right\} \\ &= \frac{m}{2}\int_{-\infty}^{+\infty}\mathrm{d}t\,j(t)\,e^{-i\omega_0 t}\int_{-\infty}^{+\infty}\mathrm{d}t'\,j(t')\,e^{+i\omega_0 t'} = \frac{m}{2}\,|\,j(\omega_0)\,|^2\,,\end{aligned} \tag{16.144}$$

which is what we were trying to show.

16.7 The three-dimensional oscillator

For a particle subject to a spherically symmetric three-dimensional harmonic restoring force, the Hamiltonian, in our units, is given by:

$$H = \frac{1}{2}\left(P_i^2 + Q_i^2\right) = N + \frac{3}{2}\,, \qquad \text{where} \qquad N = A_i^\dagger A_i\,. \tag{16.145}$$

Here the sum over i goes from 1 to 3. We note that $[\,A_i, N\,] = A_i$ and $[\,A_i^\dagger, N\,] = -A_i^\dagger$. The angular momentum operator is given by:

$$\begin{aligned} L_k &= \epsilon_{ijk}\, Q_i\, P_j = -i\,\epsilon_{ijk}\, A_i^\dagger\, A_j\,, \\ \epsilon_{ijk}\, L_k &= Q_i\, P_j - Q_j\, P_i = -i\,(\,A_i^\dagger\, A_j - A_j^\dagger\, A_i\,)\,. \end{aligned} \tag{16.146}$$

The angular momentum L_k commutes with N: $[\,L_k, N\,] = 0$ and so can be simultaniously digonalize with N. However, we generally use eigenvalues of the number operator N given by the direct product of eigenvalues of the three Cartesian number operators:

$$N\,|\,n_x, n_y, n_z\,\rangle = n\,|\,n_x, n_y, n_z\,\rangle\,, \qquad n = n_x + n_y + n_z \tag{16.147}$$

where the $n_i = 0, 1, 2, \ldots$ are non-negative integers. Here we have defined n so that n starts with zero: $n = 0, 1, 2, \cdots$. These eigenvectors are also eigenkets of the Hamiltonian, which have eigenvalues given by: $E_n = n + 3/2$. The degeneracy is given by: $(n+1)(n+2)/2$.

Exercise 45. Show that for a given n the possible values of the total angular momentum quantum number are $\ell = n, n-2, \ldots$ down to 0 or 1, and that each ℓ occurs just once.

The degeneracy of the three-dimensional harmonic oscillator indicates that something other than the angular momentum is a constant of the motion, and that there is a larger symmetry of the Hamiltonian other than $O(3)$. It is easily seen that this covering symmetry is $SU(3)$, the special group of unitary transformations in three dimensions. In fact if we consider unitary transformations of the form:

$$U^\dagger(u)\, A_i\, U(u) = u_{ij}\, A_j\,, \qquad U^\dagger(u)\, A_i^\dagger\, U(u) = u_{ij}^*\, A_j^\dagger\,, \tag{16.148}$$

where u_{ij} is a 3×3 unitary matrix, $u_{ij}^* u_{ik} = \delta_{jk}$. Then the number operator, and consequently the Hamiltonian, is invariant under this transformation:

$$U^\dagger(u)\, N\, U(u) = U^\dagger(u)\, A_i^\dagger A_i\, U(u) = u_{ij}^*\, u_{ik}\, A_j^\dagger\, A_k = A_j^\dagger\, A_j = N\,. \tag{16.149}$$

Next we find the generators of this transformation, so we put $u_{ij} = \delta_{ij} + i\Delta h_{ij} + \cdots$, where $\Delta h_{ij}^* = \Delta h_{ji}$ is a 3×3 Hermitian matrix. We also write, to first order in Δh_{ij},

$$U(1 + \Delta h) = 1 + i\Delta h_{ij}\, G_{ij} + \cdots \tag{16.150}$$

where $G_{ij}^\dagger = G_{ji}$ are generators of the infinitesimal transformation. So we find:

$$\left(1 - i\Delta h_{ij}^*\, G_{ij}^\dagger + \cdots\right) A_k \left(1 + i\Delta h_{ij}\, G_{ij} + \cdots\right) = \left(\delta_{kl} + i\,\Delta h_{kl} + \cdots\right) A_l\,,$$

or

$$A_k + i\Delta h_{ij}\,[\,A_k, G_{ij}\,] + \cdots = A_k + i\Delta h_{ij}\,\delta_{ki}\, A_j + \cdots\,,$$

from which we find:

$$[\,A_k, G_{ij}\,] = \delta_{ki}\, A_j\,, \tag{16.151}$$

the solution of which is

$$G_{ij} = A_i^\dagger\, A_j\,. \tag{16.152}$$

However, the trace of G_{ij} is just the operator N, and can be removed from the transformations by writing:

$$G_{ij} = Q_{ij} + i\epsilon_{ijk}\, L_k + \frac{1}{3}\, N\, \delta_{ij}\ , \tag{16.153}$$

where the symmetric and traceless quadrupole tensor operator Q_{ij} is defined by:

$$Q_{ij} = \frac{1}{2}\,(\, A_i^\dagger\, A_j + A_j^\dagger\, A_i\,) - \frac{1}{3}\, N\, \delta_{ij}\ , \tag{16.154}$$

and L_k is the angular momentum operator given in Eq. (16.146). Requiring the determinant of u_{ij} to be one means that we must also require the trace $\Delta h_{ii} = 0$, which eliminates the N generator. Then the five components of Q_{ij} and three components of L_k are eight generators of $SU(3)$, which is the largest symmetry group in the three-dimensional harmonic oscillator. All components of the quadrupole tensor commute with N: $[\, Q_{ij}, N\,] = 0$, as does the full generators $[\, G_{ij}, N\,] = 0$. The generators G_{ij} transform as second rank Cartesian tensors under $SU(3)$:

$$U^\dagger(u)\, G_{ij}\, U(u) = u^*_{ii'}\, u_{jj'}\, G_{i'j'}\ . \tag{16.155}$$

The angular momentum L_k transform as a pseuto-vector, or as antisymmetric components of a second rank tensor under $SU(3)$. The quadrupole tensor Q_{ij} transform as symmetric traceless components of a second rank Cartesian tensor under $SU(3)$. From Eq. (16.155), we see that the square of the generator operator G_{ij}^2 is one of the Casimir invariants:

$$U^\dagger(u)\, G_{ij}^2\, U(u) = G_{ij}^2 = Q_{ij}^2 - 2\, L_k^2 + N^2/3\,. \tag{16.156}$$

The quantum mechanical degeneracy of the three-dimensional harmonic oscillator is related to the elliptical closed orbits of the classical motion.[3]

16.8 The Fermi oscillator

In our study of identical particles, we noted that there were two types of particles found in nature, those obeying Bose statistics and those obeying Fermi statistics. Bose particles are described by operators which obey commutation relations, whereas Fermi particles obey *anti*-commutation relations. In our system of units, a Bose oscillator is described by the Hamiltonian:

$$H_B = (\, B^\dagger B + B\, B^\dagger\,)/2\,, \qquad [\, B, B^\dagger\,] = 1\,, \quad [\, B, B\,] = [\, B^\dagger, B^\dagger\,] = 0\,. \tag{16.157}$$

We define a Fermi oscillator by a similar Hamiltonian, but with operators which obey anti-commutation relations:

$$H_F = (\, F^\dagger F - F\, F^\dagger\,)/2\,, \qquad \{\, F, F^\dagger\,\} = 1\,, \quad \{\, F, F\,\} = \{\, F^\dagger, F^\dagger\,\} = 0\,. \tag{16.158}$$

The Fermi oscillator has no classical representation. The unusual anti-commutation relations for F require that:

$$F^2 = [\, F^\dagger\,]^2 = 0\,. \tag{16.159}$$

The Fermi number operator $N_F = F^\dagger F$ is Hermitian, and has a particularly simple eigenvalue spectrum, which is stated in the following theorem:

Theorem 36. *The eigenvalues and eigenvectors of the Fermi number operator are given by:*

$$N_F\,|\, n_F\,\rangle = n_F\,|\, n_F\,\rangle\,, \qquad N_F = F^\dagger F\,, \tag{16.160}$$

with $n_F = 0, 1$ *and* F *and* $F^\dagger$ *having the following action on the the two eigenvectors:*

$$F\,|\,0\,\rangle = 0\,, \qquad F\,|\,1\,\rangle = |\,0\,\rangle\,, \qquad F^\dagger\,|\,0\,\rangle = |\,1\,\rangle\,, \qquad F^\dagger\,|\,1\,\rangle = 0\,. \tag{16.161}$$

[3]For more details and references, see Schiff [**?**][pp. 234–242].

Proof. We first note that:

$$N_F^2 \,|\, n_F \,\rangle = F^\dagger F\, F^\dagger F \,|\, n_F \,\rangle = F^\dagger \,(1 - F^\dagger\, F\,)\, F \,|\, n_F \,\rangle = F^\dagger\, F \,|\, n_F \,\rangle = N_F \,|\, n_F \,\rangle \,, \tag{16.162}$$

So $n_F(n_F - 1) = 0$, which has two solutions: $n_F = 0, 1$. The rest of the proof is left to the reader. □

The energy of the Fermi oscillator is then given by:

$$H_F \,|\, n_F \,\rangle = E_{n_F} \,|\, n_F \,\rangle \,, \qquad E_{n_F} = n_F - 1/2 \,. \tag{16.163}$$

So the two energy eigenvalues of the Fermi oscillator are $\pm 1/2$. The time dependence of the Fermi operators are given by Heisenberg's equations of motion:

$$\begin{aligned} \dot{F}(t) &= [\, F(t), H \,]/i = -i\, F(t) \,, \qquad & F(t) &= F\, e^{-it} \,, \\ \dot{F}^\dagger(t) &= [\, F^\dagger(t), H \,]/i = +i\, F^\dagger(t) \,, \qquad & F^\dagger(t) &= F^\dagger\, e^{+it} \,. \end{aligned} \tag{16.164}$$

We now ask if we can find a Lagrangian for the Fermi Hamiltonian. Surprisingly we can, if we introduce Grassmann variables and a Grassmann calculus, discussed in the next section.

16.8.1 Action for a Fermi oscillator

With this introduction to Grassmann algebra, we can now write down a Lagrangian and an action for the Grassmann Fermi oscillator.[4] Let $f(t)$ and $f^*(t)$ be two complex Grassmann functions of t. Then consider the action:

$$\begin{aligned} S(f, f^*) &= \int \mathrm{d}t\, L(f, f^*; \dot{f}, \dot{f}^*) \,, \\ L(f, f^*; \dot{f}, \dot{f}^*) &= \frac{i}{2}\,(\, f^*\, \dot{f} - \dot{f}^*\, f\,) - \frac{1}{2}\,(\, f^* f - f f^*\,) \,, \end{aligned} \tag{16.165}$$

Canonical momenta and derivatives of the Lagrangian are given by:

$$\begin{aligned} p_f &= \frac{\partial L}{\partial \dot{f}} = -\frac{i}{2}\, f^* \,, \qquad & \frac{\partial L}{\partial f} &= +\frac{i}{2}\, \dot{f}^* + f^* \,, \\ p_{f^*} &= \frac{\partial L}{\partial \dot{f}^*} = -\frac{i}{2}\, f \,, \qquad & \frac{\partial L}{\partial f^*} &= +\frac{i}{2}\, \dot{f} - f \,, \end{aligned} \tag{16.166}$$

where we have used a *left* derivative convention. Lagrange's equations of motion are then:

$$\dot{f} = -if \,, \qquad \dot{f} = +if \,. \tag{16.167}$$

The Hamiltonian is given by:

$$\begin{aligned} H &= \dot{f}\, p_f + \dot{f}^* p_{f^*} - L \\ &= -\frac{i}{2}\, \dot{f}\, f^* - \frac{i}{2}\, \dot{f}^*\, f - \frac{i}{2}\,(\, f^*\, \dot{f} - \dot{f}^*\, f\,) + \frac{1}{2}\,(\, f^* f - f f^*\,) \\ &= \frac{1}{2}\,(\, f^* f - f f^*\,) \,. \end{aligned} \tag{16.168}$$

Now all four quantities, f, f^*, p_f, and p_{f^*} are *not* independent variables. Choosing f and f^* as independent variables, Hamilton's equations become:

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} f \\ f^* \end{pmatrix} = \begin{pmatrix} -if \\ if^* \end{pmatrix} = \frac{1}{i} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \partial_f H(f, f^*) \\ \partial_{f^*} H(f, f^*) \end{pmatrix} \,. \tag{16.169}$$

[4] In this section, we follow Das.

Now let A and B be any functions of the Grassmann variables (f, f^*). Then Poisson brackets for Grassmann variables are defined by:

$$\{ A, B \} = \frac{1}{i} \left(\partial_f A \,, \quad \partial_{f^*} B \right) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \partial_f A \\ \partial_{f^*} B \end{pmatrix} = \frac{1}{i} \{ \, \partial_f A \, \partial_{f^*} B + \partial_{f^*} A \, \partial_f B \, \} \,. \tag{16.170}$$

Note the plus sign and factor of $1/i$ in the definitions of the Poisson bracket for Grassmann variables. Then, for example, the classical equations of motion are given by:

$$\dot{f} = \{ \, f, H \, \} \,, \qquad \dot{f^*} = \{ \, f^*, H \, \} \,, \qquad \{ \, f, \bar{f} \, \} = 1/i \,. \tag{16.171}$$

Quantization of the classical Grassman system is carried out with the usual rules:

- Grassmann variables are mapped to Fermi operators in Hilbert space:

$$f \mapsto F \,, \qquad f^* \mapsto F^\dagger \,. \tag{16.172}$$

- Grassmann Poisson brackets are mapped to anti-commutators of the corresponding quantum operators, with a factor of $\hbar$:

$$\{ \, A, B \, \} \mapsto \{ \, A, B \, \}/i\hbar \,. \tag{16.173}$$

  In particular,

$$\{ \, F, F^\dagger \, \} = i\hbar \{ \, f, f^* \, \} = \hbar \,. \tag{16.174}$$

Since we have already used units with $\hbar = 1$, the classical Hamiltonian (16.168) becomes in quantum mechanics:

$$H = \frac{1}{2} \left(\, F^\dagger F - F \, F^\dagger \, \right) \,, \qquad \{ \, F, F^\dagger \, \} = 1 \,, \tag{16.175}$$

which is the Fermi oscillator Hamiltonian introduced in Eq. (16.158). Thus Grassmann variables has enabled us to introduce an action and a Lagrangian as in the Bose oscillator case.

References

[1] E. Schrödinger, "Der stetige Übergang von der Mikro- zur Makromechanik," Naturwissenschaften **14**, 664 (1926).

[2] P. A. M. Dirac, "La seconde quantification," Ann. Inst. H. Poincaré **11**, 15 (1949).

[3] V. Bargmann, "On a Hilbert space of analytic functions and an associated integral transform," Comm. Pure Appl. Math. **14**, 187 (1961).

[4] J. R. Klauder, "Coherent and incoherent states of the radiation field," Phys. Rev. **131**, 2766 (1963).

[5] J. R. Klauder and E. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin/Cummings Publishing, New York, NY, 1968).

[6] M. M. Nieto, "The Discovery of squeezed states — in 1927," in D. Han, J. Janszky, Y. S. Kim, and V. I. Man'ko (editors), "Fifth International Conference on Squeezed States and Uncertainty Relations," pages 175–180 (NASA, Greenbelt, MD, 200771, 1998). Quant-ph/9708012.

ANNOTATION: NASA/CP–1998–206855

[7] E. H. Kennard, "Zur Quantenmechanik einfacher Bewegungstypen," Zeit. für Physik **44**, 326 (1927).

Chapter 17

Electrons and phonons

In this chapter we develop a model for electrons and phonons on a one-dimensional lattice. We assume that the electrons experience an electrostatic potential created by the atoms making up the lattice, and that the lattice is held in place by harmonic binding forces. We consider first an approximation in which the electrons can occupy a single state at each site and can jump between neighboring sites. First order interactions between electrons and the lattice are found assuming a simple potential model.

17.1 Electron-phonon action

We describe the electrons by a non-relativistic Fermi field $\Psi(x,t)$ and the position of the $n^{\rm th}$ atom as a function of time by $\Phi_n(t)$. The displacment of the atoms from the equilibrium position $\phi_n(t)$ is defined by: $\Phi_n(t) = na + \phi_n(t)$. So our model of the classical action is given by:

$$\begin{aligned} S[\Psi,\Psi^*,\Phi] &= \int {\rm d}t\, L[\Psi,\Psi^*,\Phi;\dot\Psi,\dot\Psi^*,\dot\Phi]\,, \\ L[\Psi,\Psi^*,\Phi;\dot\Psi,\dot\Psi^*,\dot\Phi] &= L_e[\Psi,\Psi^*;\dot\Psi,\dot\Psi^*] + L_p(\Phi,\dot\Phi) + L_{ep}[\Psi,\Psi^*,\Phi]\,, \end{aligned} \tag{17.1}$$

where

$$\begin{aligned} L_e[\Psi,\Psi^*;\dot\Psi,\dot\Psi^*] &= \int_0^L {\rm d}x \left\{ \frac{i}{2}\left[\,\Psi^\dagger(x,t)\,\dot\Psi(x,t) - \dot\Psi^\dagger(x,t)\,\Psi(x,t)\,\right] - \frac{|\nabla\Psi(x,t)|^2}{2m}\right\}, \\ L_p[\Phi,\dot\Phi] &= \sum_{n=0}^{N-1}\left\{\,\frac{M}{2}\,\dot\Phi_n^2 - U(|\Phi_{n+1}(t)-\Phi_n(t)|)\,\right\}, \\ L_{ep}[\Psi,\Psi^*,\Phi] &= -\int_0^L {\rm d}x\,\sum_n V(|x-\Phi_n(t)|)\,|\Psi(x,t)|^2\,, \end{aligned} \tag{17.2}$$

where $U(\Phi_{n+1}(t)-\Phi_n(t))$ is the potential energy between atoms in the lattice and $V(x-\Phi_n(t))$ is the potential energy between the electrons and the atoms. Here m is the mass of the electrons and M the mass of the atoms.

We now write this Lagrangian to first order in the atomic displacements $\phi_n(t)$. Expanding the interatomic potential to first order, we find:

$$U(|\Phi_{n+1}(t)-\Phi_n(t)|) = U_0 + \frac{1}{2}\,M\,\omega_0^2\,[\,\phi_{n+1}(t)-\phi_n(t)\,]^2 + \cdots \tag{17.3}$$

where $M\,\omega_0^2 = [\,\partial^2 U(x)/\partial x^2\,]_{x=a}$. So the Lagrangian for the atomic motion is given immediately by:

$$L_p(\phi,\dot\phi) = \frac{M}{2}\sum_{n=0}^{N-1}\left\{\,\dot\phi_n^2(t) - \omega_0^2\,[\,\phi_{n+1}(t)-\phi_n(t)\,]^2\,\right\}. \tag{17.4}$$

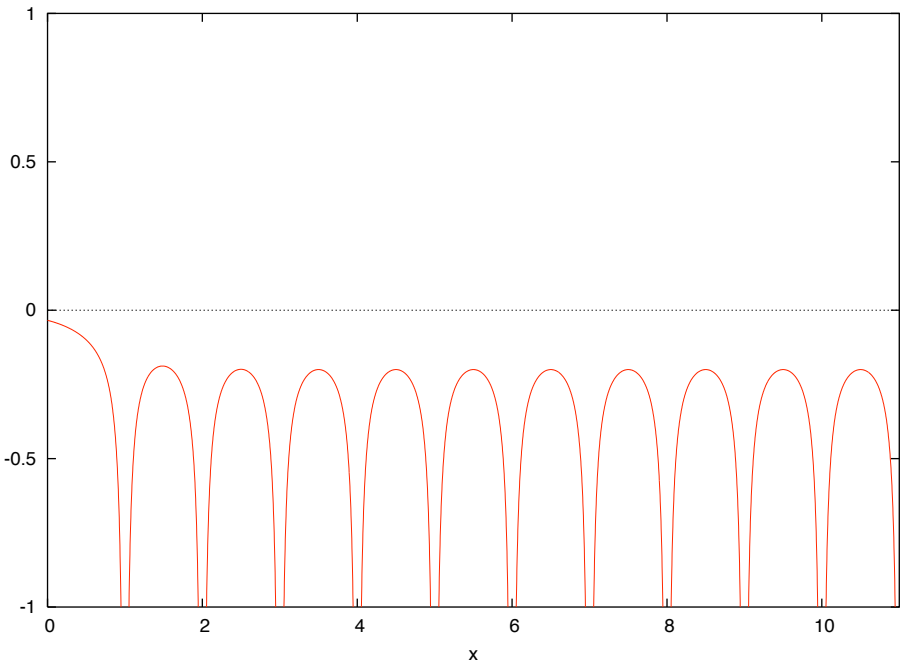

Figure 17.1: Plot of $V(x)$ for the first 10 sites.

For the electron-atom potential, we find to first order:

$$V(|x-\Phi_n(t)|) = V_n(x) + \frac{\partial V_n(x)}{\partial x}\,\phi_n(t) + \cdots\,, \qquad V_n(x) = V(|x-na|)\,. \tag{17.5}$$

Here $V(x) = \sum_n V_n(x)$ is the potential seen by the electron for the atoms located at their equilibrium positions, as shown in Fig. 17.1. So for the electron part, we introduce a set of eigenfunctions $\chi_n(x) = \psi_0(x-na)$ for the ground state of an electron located at the site n, and which satisfy:

$$h_n(x)\,\chi_n(x) = \epsilon_0\,\chi_n(x)\,, \qquad h_n(x) = -\frac{\nabla^2}{2m} + V_n(x)\,, \qquad h_0(x)\,\psi_0(x) = \epsilon_0\,\psi_0(x)\,, \tag{17.6}$$

and we expand the field $\Psi(x,t)$ in these basis functions:

$$\Psi(x,t) \approx \sum_{n=0}^{N-1} \psi_n(t)\,\chi_n(x)\,, \tag{17.7}$$

where $\psi_n(t)$ is the complex amplitude for the atom to be found in the ground state at the $n^{\rm th}$ site. In our very crude approximation here, we take the overlap integrals between neighboring sites as approximately orthogonal:

$$\int_0^L {\rm d}x\,\chi^\dagger_{n'}(x)\,\chi_n(x) \approx \delta_{n',n}\,. \tag{17.8}$$

So this means that the kinetic terms become approximately:

$$\int_0^L {\rm d}x\left\{\frac{i}{2}\left[\,\Psi^\dagger(x,t)\,\dot\Psi(x,t) - \dot\Psi^\dagger(x,t)\,\Psi(x,t)\,\right]\right. \approx \sum_{n=0}^{N-1}\frac{i}{2}\left[\,\psi^*_n\,\dot\psi_n - \dot\psi^*_n\,\psi_n\,\right]. \tag{17.9}$$

For the potential terms, we keep only the nearest neighbor interactions, so we find:

$$-\int_0^L \mathrm{d}x \sum_{n=0}^{N-1} V_n(x)\,|\Psi(x,t)|^2 \approx \sum_{n=0}^{N-1}\Big\{ -V_n(x)\,|\psi_n(t)|^2 + \Gamma\left[\,\psi^*_{n+1}(t)\psi_n(t) + \psi^*_n(t)\psi_{n+1}(t)\,\right]\Big\}$$

where

$$\Gamma = -\int \chi^*_{n+1}(x)\left[\,V_{n+1}(x) + V_n(x)\,\right]\chi_n(x)\,\mathrm{d}x > 0\,. \tag{17.10}$$

So since:

$$\int_0^L \mathrm{d}x\,\chi^\dagger_{n'}(x)\,h_n(x)\,\chi_n(x) \approx \epsilon_0\,\delta_{n'n}\,, \tag{17.11}$$

we find the Lagrangian for the electron part:

$$L_e(\psi,\psi^*;\dot\psi,\dot\psi^*) = \sum_{n=0}^{N-1}\Big\{\frac{i\hbar}{2}\left[\,\psi^*_n(t)\,\dot\psi_n(t) - \dot\psi^*_n(t)\,\psi_n(t)\,\right] - \epsilon_0\,|\,\psi_n(t)\,|^2 + \Gamma\left[\,\psi^*_{n+1}(t)\psi_n(t) + \psi^*_n(t)\psi_{n+1}(t)\,\right]\Big\}\,. \tag{17.12}$$

The electron-lattice interaction comes from the second term in Eq. (17.5). We find:

$$\begin{aligned}-\int_0^L \mathrm{d}x \sum_{n=0}^{N-1}\frac{\partial V_n(x)}{\partial x}\,\phi_n(t) &\approx -K\sum_{n=0}^{N-1}\phi_n\left[\,\psi^*_{n+1}(t)\,\psi_n(t) + \psi^*_n(t)\,\psi_{n+1}(t) - \psi^*_n(t)\,\psi_{n-1}(t) - \psi^*_{n-1}(t)\,\psi_n(t)\,\right]\\ &= -K\sum_{n=0}^{N-1}\left[\,\phi_n(t) - \phi_{n-1}(t)\,\right]\left[\,\psi^*_{n+1}(t)\,\psi_n(t) + \psi^*_n(t)\,\psi_{n+1}(t)\,\right]\,.\end{aligned} \tag{17.13}$$

where the electron-atom interaction coefficent K is written by:

$$K = \int_0^L \mathrm{d}x\,\chi^\dagger_n(x)\,\frac{\partial V_n(x)}{\partial x}\,\chi_{n+1}(x) = -\int_0^L \mathrm{d}x\,\chi^\dagger_{n-1}(x)\,\frac{\partial V_n(x)}{\partial x}\,\chi_n(x) > 0\,. \tag{17.14}$$

Then the interaction Lagrangian is given by:

$$L_{ep}(\psi,\psi^*,\phi) = -K\sum_{n=0}^{N-1}\left[\,\phi_n(t) - \phi_{n-1}(t)\,\right]\left[\,\psi^*_{n+1}(t)\,\psi_n(t) + \psi^*_n(t)\,\psi_{n+1}(t)\,\right]\,. \tag{17.15}$$

In Fig. 17.2, we show plots of $V(x)$ and $V'(x)$ for site n and the wave functions for neighboring sites. One can see from these plots that Γ and K are positive quantities.

So with these approximations, the complete Lagrangian is given by:

$$L(\psi,\psi^*,\phi;\dot\psi,\dot\psi^*,\dot\phi) = L_e(\psi,\psi^*;\dot\psi,\dot\psi^*) + L_p(\phi,\dot\phi) + L_{ep}(\psi,\psi^*,\phi) \tag{17.16}$$

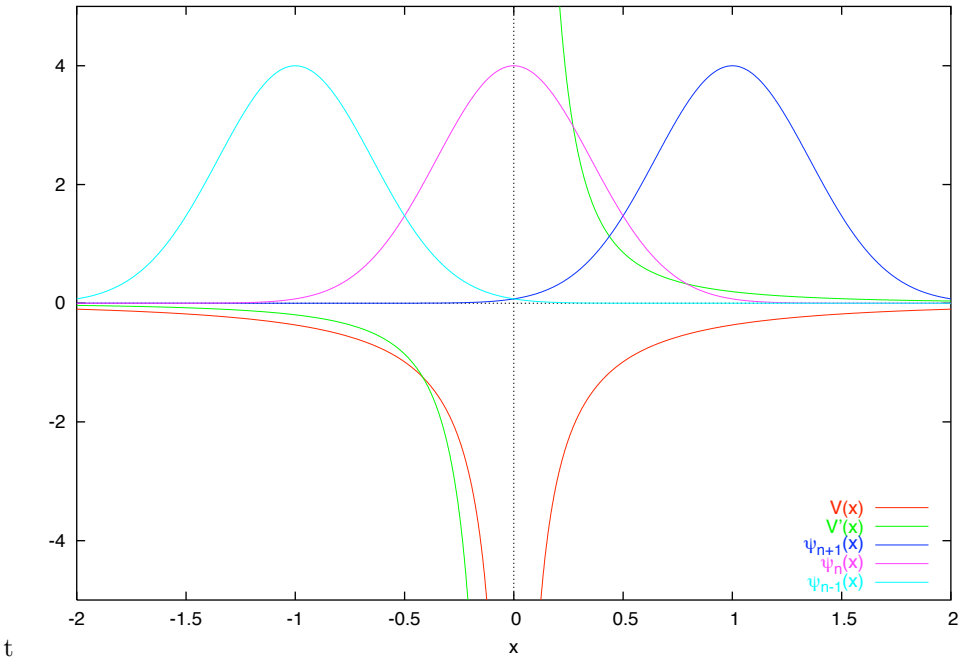

Figure 17.2: Plot of $V(x)$ and $V'(x)$ for site n with wave functions for sites n, $n \pm 1$, showing the overlap integrals between nearest neighbor sites.

where

$$L_e(\psi,\psi^*;\dot\psi,\dot\psi^*) = \sum_{n=0}^{N-1}\Big\{ \frac{i\hbar}{2}\left[\,\psi_n^*(t)\,\dot\psi_n(t) - \dot\psi_n^*(t)\,\psi_n(t)\,\right] \tag{17.17}$$
$$- \epsilon_0\,|\,\psi_n(t)\,|^2 + \Gamma\left[\,\psi_{n+1}^*(t)\psi_n(t) + \psi_n^*(t)\psi_{n+1}(t)\,\right]\Big\}\,.$$

$$L_p(\phi,\dot\phi) = \frac{M}{2}\sum_{n=0}^{N-1}\Big\{\,\dot\phi_n^2(t) - \omega_0^2\,[\,\phi_{n+1}(t)-\phi_n(t)\,]^2\,\Big\}\,. \tag{17.18}$$

$$L_{ep}(\psi,\psi^*,\phi) = -K\sum_{n=0}^{N-1}\left[\,\phi_n(t)-\phi_{n-1}(t)\,\right]\left[\,\psi_{n+1}^*(t)\,\psi_n(t)+\psi_n^*(t)\,\psi_{n+1}(t)\,\right]. \tag{17.19}$$

17.2 Equations of motion

From the Lagrangian, Eq. (17.16), we find the canonical momenta:

$$\pi_n = \frac{\partial L}{\partial\dot\psi_n} = \frac{i\hbar}{2}\,\psi_n^*\,,\qquad \pi_n^* = \frac{\partial L}{\partial\dot\psi_n^*} = -\frac{i\hbar}{2}\,\psi_n\,,\qquad p_n = \frac{\partial L}{\partial\dot\phi_n} = M\,\dot\phi_n\,, \tag{17.20}$$

and the equations of motion:

$$\begin{aligned}
i\hbar\,\dot\psi_n &= \epsilon_0\,\psi_n - \Gamma\left[\,\psi_{n+1}+\psi_{n-1}\,\right] - K\left[\,(\,\phi_{n+1}-\phi_n\,)\,\psi_{n+1} + (\,\phi_n-\phi_{n-1}\,)\,\psi_{n-1}\,\right],\\
-i\hbar\,\dot\psi_n^* &= \epsilon_0\,\psi_n^* - \Gamma\left[\,\psi_{n+1}^*+\psi_{n-1}^*\,\right] - K\left[\,(\,\phi_{n+1}-\phi_n\,)\,\psi_{n+1}^* + (\,\phi_n-\phi_{n-1}\,)\,\psi_{n-1}^*\,\right],\\
M\,\ddot\phi_n &= M\,\omega_0^2\left[\,\phi_{n+1}-2\,\phi_n+\phi_{n-1}\,\right] - K\left[\,(\,\psi_{n+1}^*-\psi_{n-1}^*\,)\,\psi_n + \psi_n^*\,(\,\psi_{n+1}-\psi_{n-1}\,)\,\right].
\end{aligned} \tag{17.21}$$

Note that the coupling between the systems is so as to provide a modified jumping probability for the electron to the next site, depending on the difference between the positions of the atoms at those sites, and that the force on an atom at a site depends on the occupation of electrons at adjacent sites. We can if we wish eliminate ϵ_0 from the dynamics by setting:

$$\psi_n(t) = \bar{\psi}_n(t)\, e^{-i\epsilon_0 t/\hbar}\,. \tag{17.22}$$

Then we find the same equations of motion as (17.21) for $\bar{\psi}_n(t)$ without the first term on the right-hand side involving ϵ_0, with a similar relation for $\bar{\psi}^*_n(t)$.

The Hamiltonian is given by a sum of three terms:

$$\begin{aligned} H(\psi,\psi^*,\phi,\dot{\phi}) &= \sum_{n=0}^{N-1}\Big\{\,\dot{\psi}_n\,\pi_n + \dot{\psi}^*_n\,\pi^*_n + \dot{\phi}_n\,p_n\,\Big\} - L(\psi,\psi^*,\phi;\dot{\psi},\dot{\psi}^*,\dot{\phi}) \\ &= H_e(\psi,\psi^*) + H_p(\phi,\dot{\phi}) + H_{ep}(\psi,\psi^*,\phi)\,, \end{aligned} \tag{17.23}$$

where

$$\begin{aligned} H_e(\psi,\psi^*) &= \sum_{n=0}^{N-1}\Big\{\,\epsilon_0\,|\,\psi_n\,|^2 - \Gamma\,\big[\,\psi^*_{n+1}\,\psi_n + \psi_{n+1}\,\psi^*_n\,\big]\,\Big\} \\ H_p(\phi,\dot{\phi}) &= \frac{M}{2}\sum_{n=0}^{N-1}\Big\{\,\dot{\phi}^2_n + \omega_0^2\,(\,\phi_{n+1}-\phi_n\,)^2\,\Big\} \\ H_{ep}(\psi,\psi^*,\phi) &= K\sum_{n=0}^{N-1}\big[\,\phi_n(t)-\phi_{n-1}(t)\,\big]\,\big[\,\psi^*_{n+1}(t)\,\psi_n(t) + \psi^*_n(t)\,\psi_{n+1}(t)\,\big]\,. \end{aligned} \tag{17.24}$$

We study the electron and photon modes for the case of no interactions between the electrons and the latice in the next two sections.

17.2.1 Numerical classical results

In this section, we describe some numerical results for the *classical* equations. We solve the classical equations of motion given in Eqs. (17.21) using a fourth-order Runga-Kutta method, as described in *Numerical Recipes* [1][p. 706]. Here, we have set $\psi_n(t) = x_n(t) + iy_n(t)$. A sample of the results are shown for the case when $\epsilon_0 = \omega_0 = \Gamma = 1$, and $K = 0.5$, for 100 sites. We show in Figs. 17.3, 17.4, 17.5, and 17.6 plots of $x_n(t)$, $y_n(t)$, $\phi_n(t)$, and $\mathrm{d}\phi_n(t)/\mathrm{d}t$, as a function of t for the first 10 sites. A movie of the same thing for all sites, as a function of time, can be found at: http://www.theory.unh.edu/graph_animation.

17.3 Electron modes

In this section, we find equations of motion for the electrons assuming no interactions with the vibrational modes of the lattice. The Lagrangian for the electrons is given by 17.12:

$$\begin{aligned} L_e(\psi,\psi^*;\dot{\psi},\dot{\psi}^*) = \sum_{n=0}^{N-1}\Big\{\,&\frac{i\hbar}{2}\,\big[\,\psi^*_n(t)\,\dot{\psi}_n(t) - \dot{\psi}^*_n(t)\,\psi_n(t)\,\big] \\ &- \epsilon_0\,|\,\psi_n(t)\,|^2 + \Gamma\,\big[\,\psi^*_{n+1}(t)\psi_n(t) + \psi^*_n(t)\psi_{n+1}(t)\,\big]\,\Big\}\,. \end{aligned} \tag{17.25}$$

The boundary conditions require that $\psi_N(t) = \psi_0(t)$. The canonical momentua is:

$$\pi_n = \frac{\partial L}{\partial\dot{\psi}_n} = \frac{i\hbar}{2}\,\psi^*_n\,, \qquad \pi^*_n = \frac{\partial L}{\partial\dot{\psi}^*_n} = -\frac{i\hbar}{2}\,\psi_n\,, \tag{17.26}$$

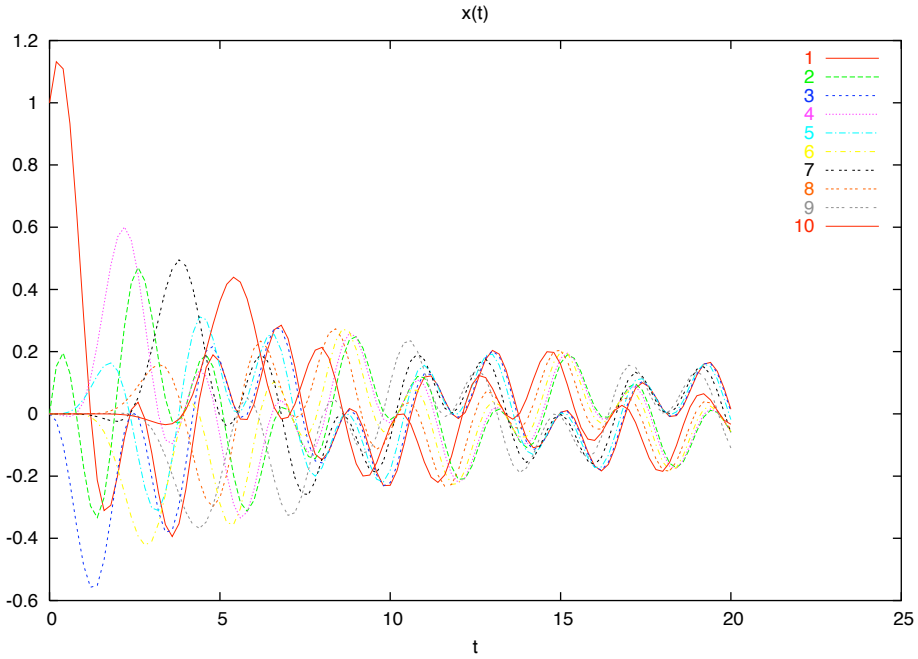

Figure 17.3: Plot of $x_n(t)$ for the first 10 sites as a function of time for $\epsilon_0 = \omega_0 = \Gamma = 1$, and $K = 0.5$, for 100 sites.

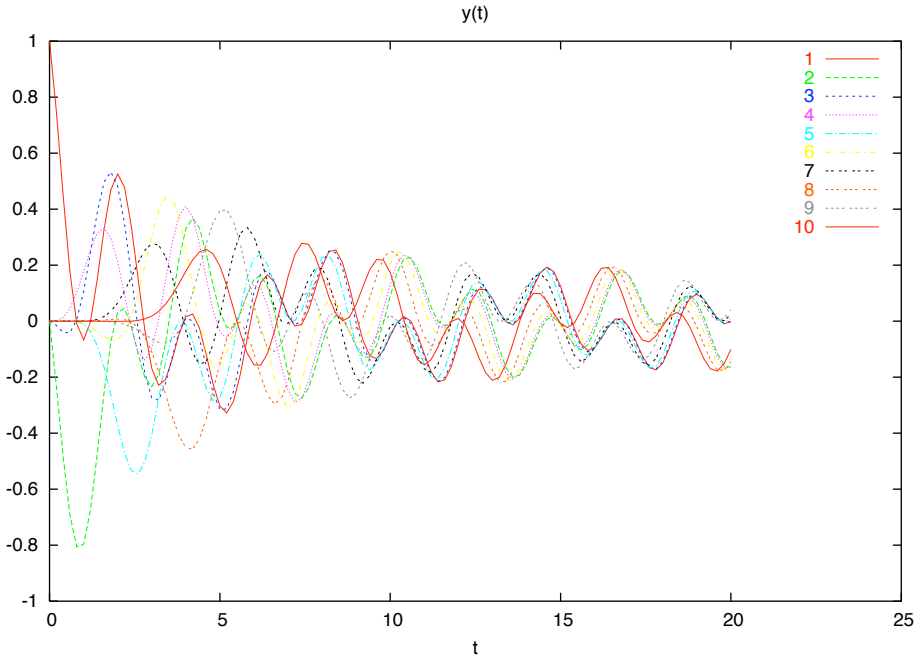

Figure 17.4: Plot of $y_n(t)$ for the first 10 sites as a function of time for $\epsilon_0 = \omega_0 = \Gamma = 1$, and $K = 0.5$, for 100 sites.

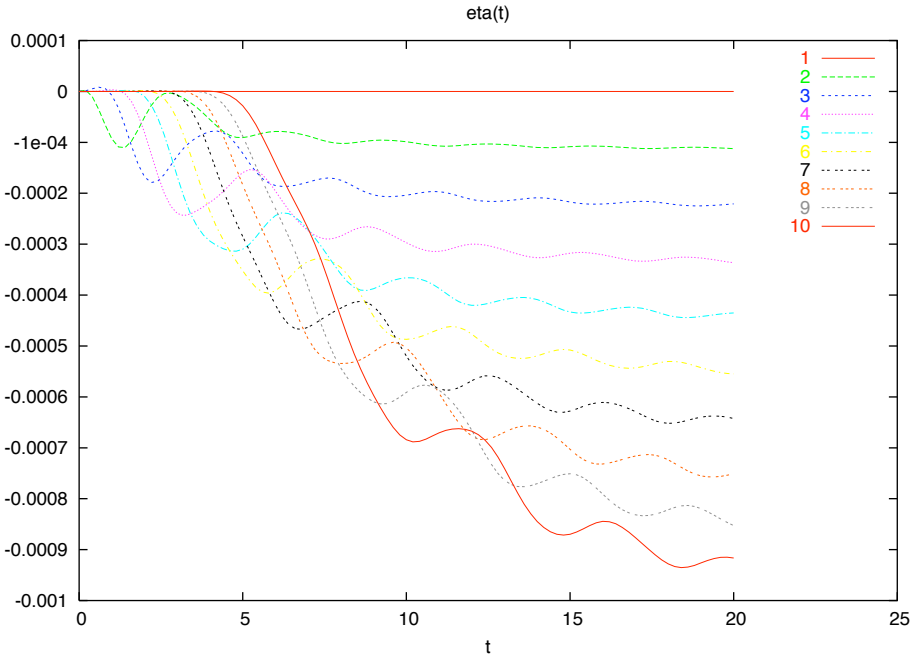

Figure 17.5: Plot of $\phi_n(t)$ for the first 10 sites as a function of time for $\epsilon_0 = \omega_0 = \Gamma = 1$, and $K = 0.5$, for 100 sites.

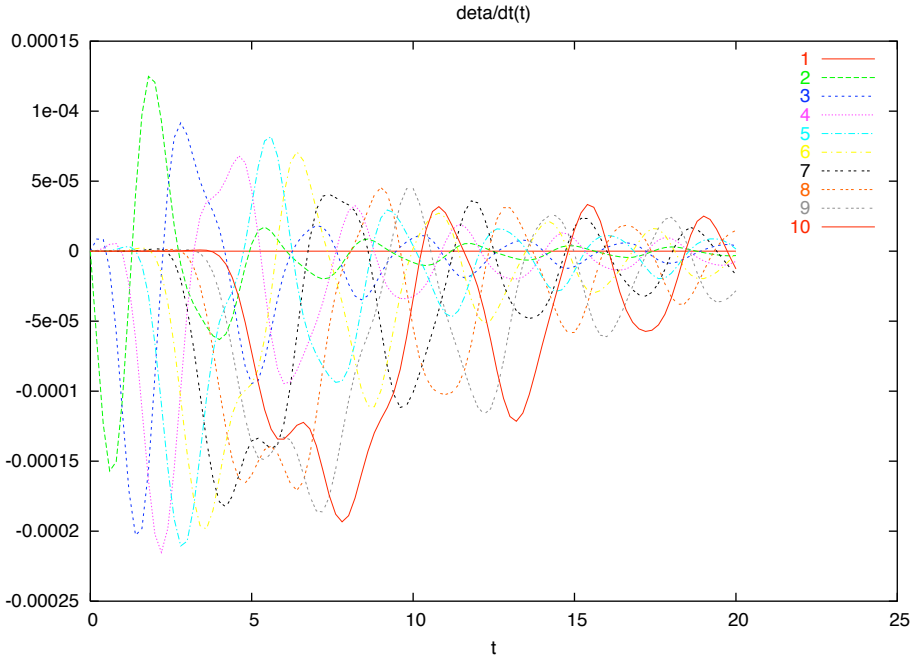

Figure 17.6: Plot of $\mathrm{d}\phi_n(t)/\mathrm{d}t$ for the first 10 sites as a function of time for $\epsilon_0 = \omega_0 = \Gamma = 1$, and $K = 0.5$, for 100 sites.

and the equations of motion are:

$$\begin{aligned} i\hbar\,\dot{\psi}_n &= \epsilon_0\,\psi_n - \Gamma\left[\,\psi_{n+1} + \psi_{n-1}\,\right], \\ -i\hbar\,\dot{\psi}^*_n &= \epsilon_0\,\psi^*_n - \Gamma\left[\,\psi^*_{n+1} + \psi^*_{n-1}\,\right], \end{aligned} \tag{17.27}$$

and the Hamiltonian is:

$$\begin{aligned} H(\psi,\psi^*) &= \sum_{n=0}^{N-1}\left[\,\dot{\psi}_n\,\pi_n + \dot{\psi}^*_n\,\pi^*_n\,\right] - L(\psi,\psi^*;\dot{\psi},\dot{\psi}^*) \\ &= \sum_{n=0}^{N-1}\Big\{\,\epsilon_0\,|\,\psi_n(t)\,|^2 - \Gamma\left[\,\psi^*_{n+1}(t)\psi_n(t) + \psi^*_n(t)\psi_{n+1}(t)\,\right]\Big\}. \end{aligned} \tag{17.28}$$

We have previously explained in Section 16.3 that the sympletic variables here are ψ_n and ψ^*_n. This form of the Hamiltonian is the same as what we had for the $N = 2$ case of a diatomic molecule, only now written in terms the amplitudes $\psi_n(t)$ for finding the electron at N atomic sites. Wave functions for electrons are described by the sympletic anticommuting operators $\psi(t)$ and $\psi^*(t)$ which satisfy the anticommutation relations (See Sec. **??**):

$$\{\,\psi_n(t),\psi^\dagger_{n'}(t)\,\} = \delta_{n,n'}\,. \tag{17.29}$$

The periodic requirement can be satisfied by finding solutions in the form of finite Fourier transforms. We put:

$$\begin{aligned} \psi_n(t) &= \frac{1}{\sqrt{N}}\sum_{k=-[N/2]+1}^{[N/2]}\tilde{\psi}_k(t)\,e^{+i\theta_k\,n}\,, \\ \psi^\dagger_n(t) &= \frac{1}{\sqrt{N}}\sum_{k=-[N/2]+1}^{[N/2]}\tilde{\psi}^\dagger_k(t)\,e^{-i\theta_k\,n}\,, \end{aligned} \tag{17.30}$$

where $\theta_k N = 2\pi k$. Then $\psi_0(t) = \psi_N(t)$. The inverse relations are:

$$\begin{aligned} \tilde{\psi}_k(t) &= \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\psi_n(t)\,e^{-i\theta_k\,n}\,, \\ \tilde{\psi}^\dagger_k(t) &= \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\psi^\dagger_n(t)\,e^{+i\theta_k\,n}\,, \end{aligned} \tag{17.31}$$

The finite Fourier transforms then obey the algebra:

$$\begin{aligned} \{\,\tilde{\psi}_k(t),\tilde{\psi}^\dagger_{k'}(t)\,\} &= \frac{1}{N}\sum_{n,n'=0}^{N-1}\{\,\psi_n(t),\psi^\dagger_{n'}(t)\,\}e^{+i(\,\theta_{k'}\,n'-\theta_k n\,)} \\ &= \frac{1}{N}\sum_{n=0}^{N-1}e^{+i(\,\theta_{k'}-\theta_k\,)n} = \delta_{k,k'}\,. \end{aligned} \tag{17.32}$$

The equations of motion then become:

$$\begin{aligned} i\hbar\,\dot{\tilde{\psi}}_k(t) &= \left[\,\epsilon_0 - 2\Gamma\,\cos(\theta_k)\,\right]\tilde{\psi}_k(t) \equiv \epsilon_k\,\tilde{\psi}_k(t)\,, \\ -i\hbar\,\dot{\tilde{\psi}}^\dagger_k(t) &= \left[\,\epsilon_0 - 2\Gamma\,\cos(\theta_k)\,\right]\tilde{\psi}^\dagger_k(t) \equiv \epsilon_k\,\tilde{\psi}^\dagger_k(t)\,, \end{aligned} \tag{17.33}$$

which have solutions:

$$\tilde{\psi}_k(t) = c_k\,e^{-i\epsilon_k t/\hbar}\,, \qquad \tilde{\psi}^\dagger_k(t) = c^\dagger_k\,e^{+i\epsilon_k t/\hbar}\,, \tag{17.34}$$

where

$$\epsilon_k = \epsilon_0 - 2\Gamma\,\cos(2\pi k/N)\,. \tag{17.35}$$

Note that for k in the range $-[N/2]-1 \le k \le [N/2]$, $\epsilon_{-k} = \epsilon_k$. There are exactly N eigenvalues. If we let a be the distance between atomic sites, then it is useful to define:

$$x_n = n\,a\,, \qquad p_k = \frac{\theta_k}{a} = \frac{2\pi\hbar\,k}{L}\,, \qquad \text{so that:} \qquad \theta_k\,n = p_k\,x_n/\hbar\,. \tag{17.36}$$

where $L = aN$ is the length around the circular chain. The operator amplitudes for finding the electron at site n are then given by:

$$\begin{aligned}\psi_n(t) &= \frac{1}{\sqrt{N}}\sum_{k=-[N/2]+1}^{[N/2]} c_k\,e^{+i(\,p_k x_n - \epsilon_k t\,)/\hbar}\,,\\ \psi_n^\dagger(t) &= \frac{1}{\sqrt{N}}\sum_{k=-[N/2]+1}^{[N/2]} c_k^\dagger\,e^{+i(\,p_k x_n - \epsilon_k t\,)/\hbar}\,.\end{aligned} \tag{17.37}$$

where the time independent operators c_k and $c_k^\dagger$ obey the anticommutator algebra:

$$\{\,c_k, c_{k'}^\dagger\,\} = \delta_{k,k'}\,. \tag{17.38}$$

Using these solutions, the Hamiltonian is given by:

$$H = \frac{1}{2}\sum_{k=-[N/2]+1}^{[N/2]} \epsilon_k\left\{\,c_k^\dagger\,c_k - c_k\,c_k^\dagger\,\right\}. \tag{17.39}$$

Each mode k has a number operator $N_k = c_k^\dagger\,c_k$ which is Hermitian and has two eigenvalues:

$$N_k\,|\,n_k\,\rangle = n_k\,|\,n_k\,\rangle\,, \qquad \text{where} \qquad n_k = 0, 1. \tag{17.40}$$

A basis set for the system is then given by the direct product of the eigenvectors for each k mode. We let n be the set of integers having integer values of zero or one for each mode: $n = \{\,n_0, n_{\pm1}, n_{\pm2}, \ldots, n_{+[N/2]}\,\}$, and write the eigenvector for this set in a short-hand notation:

$$|\,n\,\rangle \equiv |\,n_0, n_{\pm1}, n_{\pm2}, \ldots, n_{+[N/2]}\,\rangle\,. \tag{17.41}$$

The Hamiltonian is diagonal in these eigenvectors:

$$H\,|\,n\,\rangle = E_n\,|\,n\,\rangle\,, \qquad E_n = \sum_{k=-[N/2]+1}^{[N/2]} \epsilon_k\left\{\,n_k - \frac{1}{2}\,\right\}. \tag{17.42}$$

The excitation of a single k mode produces a travelling electron wave on the lattice so that the electron is distributed over the entire lattice. The dispersion of the wave is shown in Fig. 17.7. For large N and small value of k, we find that:

$$\epsilon_k \approx E_0 + \frac{p_k^2}{2m^*} + \cdots\,, \tag{17.43}$$

where $E_0 = \epsilon_0 - 2\Gamma$, and the effective mass m^* is given by:

$$m^* = \frac{\hbar^2}{2\,\Gamma\,a^2} > 0\,, \tag{17.44}$$

which has nothing to do with the physical mass, but with the transition rate between adjacent atomic sites. Because of the Pauli principle and the spin of the electron, at zero temperature the electrons fill up the available states to a final state labeled by k_F.

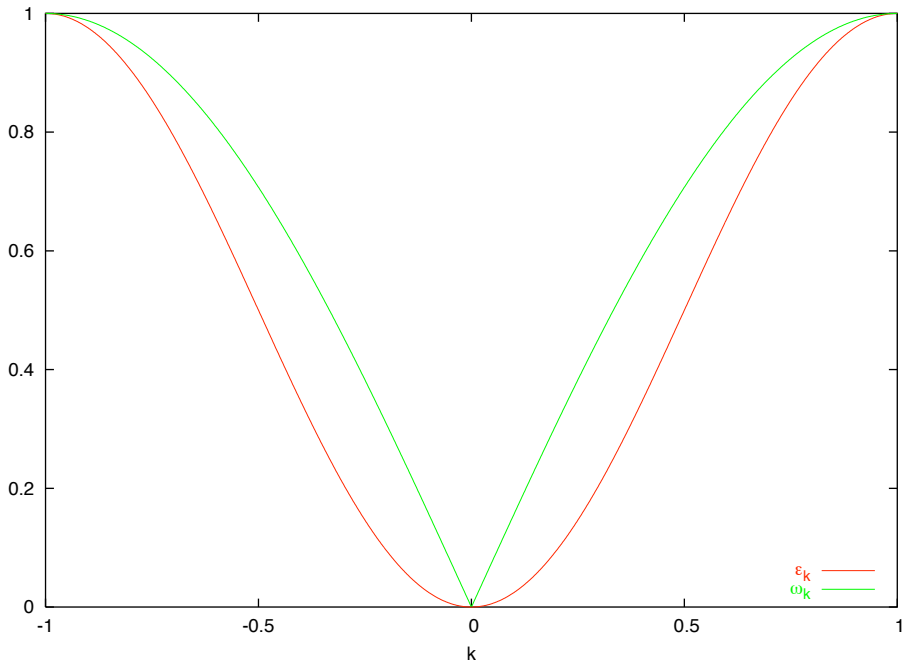

Figure 17.7: Plot of the electron and phonon energy spectra ϵ_k and ω_k on the periodic chain, as a function of k. Energies and k values have been normalized to unity. Note that near $k = 0$, the electron spectra is quadratic whereas the phonon spectrum is linear.

17.4 Vibrational modes

For the vibrational modes, the Lagrangian is given by (17.4):

$$L_p(\phi,\dot{\phi}) = \frac{M}{2}\sum_{n=0}^{N-1}\left\{\,\dot{\phi}_n^2(t) - \omega_0^2\,[\,\phi_{n+1}(t) - \phi_n(t)\,]^2\,\right\}. \tag{17.45}$$

The periodic condition requires that $\phi_0 = \phi_N$. The canonical momenta is:

$$\pi_n = \frac{\partial L}{\partial\dot{\phi}_n} = m\,\dot{\phi}_n\,, \tag{17.46}$$

and the equations of motion are:

$$\ddot{\phi}_n - \omega_0^2\,(\,\phi_{n+1} - 2\,\phi_n + \phi_{n-1}\,) = 0\,, \tag{17.47}$$

and the Hamiltonian is given by:

$$H(\phi,\pi) = \sum_{n=0}^{N-1}\left\{\,\frac{\pi_n^2}{2M} + \frac{1}{2}\,M\omega_0^2\,(\,\phi_{n+1} - \phi_n\,)^2\,\right\}. \tag{17.48}$$

The equations of motion can be solved by introducing solutions of the form:

$$\phi_n(t) = \tilde{\phi}_k(t)\,e^{i\theta_k n}\,. \tag{17.49}$$

The periodic condition requires that θ_k satisfy $\theta_k N = 2\pi k$, with k an integer. Then Eq. (17.47) becomes:

$$\ddot{\tilde{\phi}}_k(t) + \omega_k^2\,\tilde{\phi}_k(t) = 0\,. \tag{17.50}$$

where

$$\omega_k^2 = 2\omega_0^2\,(\,1-\cos(\theta_k)\,) = 4\omega_0^2\,\sin^2(\theta_k/2)\,. \tag{17.51}$$

choosing $\omega_{|k|}$ to be the *positive* root:

$$\omega_{|k|} = 2\omega_0\sin(\theta_{|k|}/2) = 2\omega_0\sin(\,\pi|k|/N\,) > 0\,, \tag{17.52}$$

the *classical* solutions of (17.50) can be written in the form:

$$\tilde{\phi}_k(t) = \tilde{a}_k\,e^{-i\omega_{|k|}t} + \tilde{a}_k^*\,e^{+i\omega_{|k|}t}\,. \tag{17.53}$$

The linear combination of these solutions for all values of k give the general classical solution to the vibrations of the molecule.

It will be useful for the quantum problem to define p_k by:

$$p_k = \frac{\hbar\,\theta_k}{a} = \frac{2\pi\hbar\,k}{L}\,, \qquad \text{so that:} \qquad \theta_k\,n = p_k x_n/\hbar\,, \tag{17.54}$$

with $L = aN$. Canonical quantization of the vibrations requires that ϕ_n and π_n become operators satisfying the algebra:

$$[\,\phi_n(t),\pi_{n'}(t)\,] = i\hbar\,\delta_{n,n'}\,. \tag{17.55}$$

As we have learned, because of the periodic condition, it is useful to introduce finite Fourier transforms:

$$\begin{aligned}\phi_n(t) &= \frac{1}{\sqrt{N}}\sum_{k=-[N/2]+1}^{[N/2]}\tilde{\phi}_k(t)\,e^{ip_kx_n/\hbar}\,,\\ \pi_n(t) &= \frac{1}{\sqrt{N}}\sum_{k=-[N/2]+1}^{[N/2]}\tilde{\pi}_k(t)\,e^{ip_kx_n/\hbar}\,,\end{aligned} \tag{17.56}$$

where $\tilde{\pi}_k(t) = m\,\dot{\tilde{\phi}}_k(t)$. Here we have introduced a factor of $1/\sqrt{N}$ in our definitions. Since $\phi_n(t)$ and $\pi_n(t)$ are real,

$$\tilde{\phi}_k^*(t) = \tilde{\phi}_{-k}(t)\,, \qquad \tilde{\pi}_k^*(t) = \tilde{\pi}_{-k}(t)\,. \tag{17.57}$$

The inverse relations are:

$$\begin{aligned}\tilde{\phi}_k(t) &= \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\phi_n(t)\,e^{-ip_kx_n/\hbar}\,,\\ \tilde{\pi}_k(t) &= \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\pi_n(t)\,e^{-ip_kx_n/\hbar}\,.\end{aligned} \tag{17.58}$$

So the commutation relations for the finite Fourier transforms are:

$$\begin{aligned}[\,\tilde{\phi}_k(t),\tilde{\pi}_{k'}^\dagger(t)\,] &= \frac{1}{N}\sum_{n,n'=0}^{N-1}[\,\phi_n(t),\pi_{n'}(t)\,]e^{i(p_{k'}x_{n'}-p_kx_n)/\hbar}\\ &= \frac{i\hbar}{N}\sum_{n=0}^{N-1}e^{i(p_{k'}-p_k)\,x_n/\hbar} = i\hbar\,\delta_{k,k'}\,.\end{aligned} \tag{17.59}$$

Using the orthorgonality realationships for finite Fourier transforms, the kintic energy becomes:

$$\begin{aligned}T &= \frac{1}{2M}\sum_{n=0}^{N-1}\pi_n^2(t) = \frac{1}{2MN}\sum_{k,k'}\tilde{\pi}_{k'}^*(t)\tilde{\pi}_k(t)\sum_{n=0}^{N-1}e^{i(p_k-p_{k'})\,x_n/\hbar}\\ &= \frac{1}{2M}\sum_{k=-[N/2]+1}^{[N/2]}\bigl|\,\pi_k(t)\,\bigr|^2 = \frac{M}{2}\sum_{k=-[N/2]+1}^{[N/2]}\bigl|\,\dot{\tilde{\phi}}_k(t)\,\bigr|^2\,,\end{aligned}$$

and for the potential part, we find:

$$\begin{aligned} V &= \frac{1}{2}\,M\omega_0^2 \sum_{n=0}^{N-1} (\,\phi_{n+1}(t) - \phi_n(t)\,)^2 \\ &= \frac{M\omega_0^2}{2N} \sum_{k,k'} \tilde{\phi}^*_{k'}(t)\tilde{\phi}_k(t)(\,e^{+i\theta_k} - 1\,)\,(\,e^{-i\theta_{k'}} - 1\,) \sum_{n=1}^{N} e^{i(p_k - p_{k'})\,x_n} \\ &= \frac{M}{2} \sum_{k=-[N/2]+1}^{[N/2]} \omega^2_{|k|}\,\big|\,\tilde{\phi}_k(t)\,\big|^2\,. \end{aligned}$$

So the Lagrangian and Hamiltonian can be written as:

$$L = \frac{M}{2} \sum_{k=-[N/2]+1}^{[N/2]} \left\{\,\big|\,\dot{\tilde{\phi}}_k(t)\,\big|^2 - \omega^2_{|k|}\,\big|\,\tilde{\phi}_k(t)\,\big|^2\,\right\}, \tag{17.60}$$

$$H = \frac{M}{2} \sum_{k=-[N/2]+1}^{[N/2]} \left\{\,\big|\,\dot{\tilde{\phi}}_k(t)\,\big|^2 + \omega^2_{|k|}\,\big|\,\tilde{\phi}_k(t)\,\big|^2\,\right\}. \tag{17.61}$$

Introducing the non-Hermitian operators $a_k(t)$ and $a^\dagger_k(t)$:

$$\begin{aligned} \tilde{\phi}_k(t) &= \sqrt{\frac{\hbar}{2M\omega_{|k|}}}\,\left[\,a_k(t) + a^\dagger_{-k}(t)\,\right], \qquad a_k(t) = \sqrt{\frac{M\omega_{|k|}}{2\hbar}}\,\tilde{\phi}_k(t) + \frac{1}{i}\,\sqrt{\frac{1}{2M\hbar\omega_{|k|}}}\,\tilde{\pi}_k(t)\,, \\ \tilde{\pi}_k(t) &= i\sqrt{\frac{M\hbar\omega_{|k|}}{2}}\,\left[\,a_k(t) - a^\dagger_{-k}(t)\,\right], \qquad a^\dagger_{-k}(t) = \sqrt{\frac{M\omega_{|k|}}{2\hbar}}\,\tilde{\phi}_k(t) - \frac{1}{i}\,\sqrt{\frac{1}{2M\hbar\omega_{|k|}}}\,\tilde{\pi}_k(t)\,. \end{aligned} \tag{17.62}$$

from which we find:

$$[\,a_k(t), a^\dagger_{k'}(t)\,] = \delta_{k,k'}\,, \tag{17.63}$$

with all other commutators vanishing. In terms of these variables, the Lagrangian and Hamiltonian becomes:

$$\begin{aligned} L &= \frac{-1}{2} \sum_{k=-[N/2]+1}^{[N/2]} \hbar\omega_{|k|}\left\{\,a^\dagger_k(t)\,a^\dagger_{-k}(t) + a_k(t)\,a_{-k}(t)\,\right\}, \\ H &= \frac{1}{2} \sum_{k=-[N/2]+1}^{[N/2]} \hbar\omega_{|k|}\left\{\,a^\dagger_k(t)\,a_k(t) + a_k(t)\,a^\dagger_k(t)\,\right\}. \end{aligned} \tag{17.64}$$

In terms of these variables, the displacement and canonical momentum is given by:

$$\begin{aligned} \phi_n(t) &= \sum_{k=-[N/2]+1}^{[N/2]} \sqrt{\frac{\hbar}{2NM\omega_{|k|}}}\left\{\,a_k(t) + a^\dagger_{-k}(t)\,\right\} e^{ip_k x_n/\hbar}\,, \\ &= \sum_{k=-[N/2]+1}^{[N/2]} \sqrt{\frac{\hbar}{2NM\omega_{|k|}}}\left\{\,a_k(t)\,e^{+ip_k x_n/\hbar} + a^\dagger_k(t)\,e^{-ip_k x_n/\hbar}\,\right\} \\ \pi_n(t) &= i \sum_{k=-[N/2]+1}^{[N/2]} \sqrt{\frac{M\hbar\omega_{|k|}}{2N}}\left\{\,a_k(t) - a^\dagger_{-k}(t)\,\right\} e^{ip_k x_n/\hbar}\,, \\ &= i \sum_{k=-[N/2]+1}^{[N/2]} \sqrt{\frac{M\hbar\omega_{|k|}}{2N}}\left\{\,a_k(t)\,e^{+ip_k x_n/\hbar} - a^\dagger_k(t)\,e^{-ip_k x_n/\hbar}\,\right\} \end{aligned} \tag{17.65}$$

Each mode k has a number operator $N_k = a_k^\dagger a_k$ which is Hermitian and has non-negative integers as eigenvalues:

$$N_k \,|\, n_k \,\rangle = n_k \,|\, n_k \,\rangle \,, \qquad \text{where} \qquad n_k = 0, 1, 2, \ldots \,. \tag{17.66}$$

A basis set for the system is then given by the direct product of the eigenvectors for each k mode. We let n be the set of integers for each mode: $n = \{\, n_{\pm 1}, n_{\pm 2}, \ldots, n_{\pm [N/2]} \,\}$, and write the eigenvector for this set in a short-hand notation:

$$|\, n \,\rangle \equiv |\, n_{\pm 1}, n_{\pm 2}, \ldots, n_{\pm [N/2]} \,\rangle \,. \tag{17.67}$$

Here we have omitted the spurious $k = 0$ mode, which represents a translation of the system, and will be discussed below. The Hamiltonian is diagonal in these eigenvectors:

$$H \,|\, n \,\rangle = E_n \,|\, n \,\rangle \,, \qquad E_n = \sum_{k=-[N/2]+1}^{[N/2]} \hbar\omega_{|k|} \left\{ n_k + \frac{1}{2} \right\} . \tag{17.68}$$

This represents harmonic oscillator vibrations built on each mode k. The dynamics is best solved using Heisenberg's equations of motion for $a_k(t)$. We have:

$$\frac{\mathrm{d}a_k(t)}{\mathrm{d}t} = \frac{[\, a_k(t), H \,]}{i\hbar} = -i\omega_{|k|}\, a_k(t) \,, \tag{17.69}$$

so

$$a_k(t) = a_k\, e^{-i\omega_{|k|}t} \,, \qquad a_k^\dagger(t) = a_k^\dagger\, e^{+i\omega_{|k|}t} \,. \tag{17.70}$$

where a_k and $a_k^\dagger$ are constant operators. The displacement operators are then given by:

$$\phi_n(t) = \sum_{k=-[N/2]+1}^{[N/2]} \sqrt{\frac{\hbar}{2NM\omega_{|k|}}} \left\{ a_k\, e^{+i(p_k x_n - e_k t)/\hbar} + a_k^\dagger\, e^{-i(p_k x_n - e_k t)/\hbar} \right\} , \tag{17.71}$$

where $e_k = \hbar\omega_{|k|}$. The excitation of a single k mode, which involve a travelling compressional wave on the lattice, is called a **phonon**. The dispersion of the wave is shown in Fig. 17.7. For large values of N and small values of k, we find that:

$$e_k = \hbar\omega_k \approx a\omega_0 \frac{2\pi\hbar|k|}{L} = v_0\, p_{|k|} \,, \tag{17.72}$$

where $v_0 = a\omega_0$ is the group velocity. Thus for small k, the wave travels without dispersion.

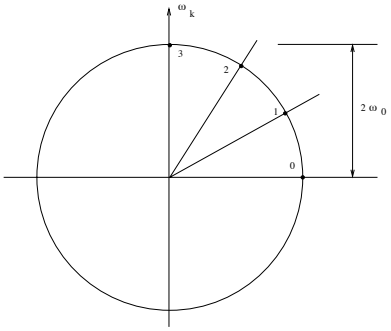

Figure 17.8: Construction for finding the oscillation frequencies for the six periodic sites of Fig. 12.3, for values of $k = 0, \pm 1, \pm 2, +3$.

Exercise 46. Show that the six vibrational frequencies of the $N = 6$ periodic chain of molecules shown in Fig. 12.3 is given by the construction shown in Fig. 17.8.

17.5 Electron-phonon interaction

From Eq. (17.24), we find:

$$H_{ep}(\psi,\psi^*,\phi) = K\sum_{n=0}^{N-1}\left[\,\phi_n(t)-\phi_{n-1}(t)\,\right]\left[\,\psi^*_{n+1}(t)\,\psi_n(t)+\psi^*_n(t)\,\psi_{n+1}(t)\,\right]. \tag{17.73}$$

Using the finite Fourier mode expansions given in Eqs. (17.30) and (17.56) for the electron and phonon modes, we find:

$$\psi^*_{n+1}(t)\,\psi_n(t) = \frac{1}{N^2}\sum_{k,k'=-[N/2]}^{[N/2]}\tilde\psi^*_{k'}(t)\,\tilde\psi_k(t)\,e^{i2\pi n(k-k')/N}\,e^{-i2\pi k'/N}\,,$$

$$\psi^*_n(t)\,\psi_{n+1}(t) = \frac{1}{N^2}\sum_{k,k'=-[N/2]}^{[N/2]}\tilde\psi^*_{k'}(t)\,\tilde\psi_k(t)\,e^{i2\pi n(k-k')/N}\,e^{+i2\pi k/N}\,,$$

$$\phi_n(t)-\phi_{n-1}(t) = \frac{1}{N}\sum_{q=-[N/2]}^{[N/2]}\tilde\phi_q(t)\left[\,1-e^{-i2\pi q/N}\,\right]e^{+i2\pi nq/N}\,.$$

So using orthogonality relations, H_{ep} becomes:

$$H_{ep}(\psi,\psi^*,\phi) = \frac{1}{N^2}\sum_{k,k'=-[N/2]}^{[N/2]}V_{k',k}\,\tilde\phi_{k'-k}(t)\,\tilde\psi^*_{k'}(t)\,\tilde\psi_k(t)\,, \tag{17.74}$$

where

$$\begin{aligned} V_{k',k} &= 2\,i\,K\left[\,\sin(2\pi k'/N)-\sin(2\pi k/N)\,\right]\\ &= 4\,i\,K\,\sin(\pi(k'-k)/N)\cos(\pi(k'+k)/N)\\ &\approx 4\pi\,i\,K\,(k'-k)/N\,. \end{aligned} \tag{17.75}$$

The last line is valid only for small values of k and k'. Note that $V^*_{k',k} = V_{k,k'} = -V_{k',k}$. From Eq. (17.62), we can also write this as:

$$H_{ep}(\psi,\psi^*,\phi) = \frac{1}{N^2}\sum_{k,k'=-[N/2]}^{[N/2]}M_{k',k}\left[\,a_{k'-k}(t)+a^*_{k-k'}(t)\,\right]\tilde\psi^*_{k'}(t)\,\tilde\psi_k(t)\,, \tag{17.76}$$

where

$$M_{k',k} = \sqrt{\frac{\hbar}{2M\omega_{|k'-k|}}}\,V_{k',k}\,. \tag{17.77}$$

So the normal mode expansion of the Hamiltonian is given by:

$$\begin{aligned} H = \frac{1}{N}\sum_{k=-[N/2]}^{[N/2]}&\left\{\,\epsilon_k\,\tilde\psi^*_k(t)\,\tilde\psi_k(t)+\omega_k\,a^*_k(t)\,a_k(t)\,\right\}\\ &+\frac{1}{N^2}\sum_{k,k'=-[N/2]}^{[N/2]}M_{k',k}\left\{\,a_{k'-k}(t)\,+a^*_{k-k'}(t)\,\right\}\tilde\psi^*_{k'}(t)\,\tilde\psi_k(t)\,. \end{aligned} \tag{17.78}$$

Eq. (17.78) is called the **Fröhlich** Hamiltonian.

17.6 The action revisited

We can write the action as:

$$S[\tilde{\psi},\tilde{\psi}^*,\phi] = \int \mathrm{d}t\, L(\tilde{\psi},\tilde{\psi}^*,\phi,\dot{\phi})\,, \tag{17.79}$$

where the Lagrangian is now given by the normal mode expansion:

$$L(\tilde{\psi},\tilde{\psi}^*,\phi,\dot{\phi}) = \frac{1}{N}\sum_{k=-[N/2]}^{[N/2]}\left\{\frac{i}{2}\left[\,\tilde{\psi}_k^*(t)\,\dot{\tilde{\psi}}_k(t) - \dot{\tilde{\psi}}_k^*(t)\,\tilde{\psi}_k(t)\,\right] - \epsilon_k\,|\tilde{\psi}_k(t)|^2\right.$$
$$\left. + \frac{m}{2}\left[\,|\dot{\tilde{\phi}}_k(t)|^2 - \omega_k^2\,|\tilde{\phi}_k(t)|^2\,\right]\right\} - \frac{1}{N^2}\sum_{k,k'=-[N/2]}^{[N/2]} V_{k',k}\,\tilde{\phi}_{k'-k}(t)\,\tilde{\psi}_{k'}^*(t)\,\tilde{\psi}_k(t)\,. \tag{17.80}$$

Here $\tilde{\psi}_k(t)$ and $\tilde{\psi}_k^*(t)$ are Grassmann variables. $\phi_k(t)$ is an ordinary commuting variable. We can write the action in a more compact form by integrating by parts over t. We first introduce two vectors:

$$\phi_k(t) = \begin{pmatrix}\tilde{\phi}_k(t)\\ \tilde{\phi}_k^*(t)\end{pmatrix}\,,\qquad \phi_k^\dagger(t) = \begin{pmatrix}\tilde{\phi}_k^*(t), & \tilde{\phi}_k(t)\end{pmatrix}\,, \tag{17.81}$$

and

$$\chi_k(t) = \begin{pmatrix}\tilde{\psi}_k(t)\\ \tilde{\psi}_k^*(t)\end{pmatrix}\,,\qquad \chi_k^\dagger(t) = \begin{pmatrix}\tilde{\psi}_k^*(t), & \tilde{\psi}_k(t)\end{pmatrix}\,, \tag{17.82}$$

and the inverse Green function operators:

$$G^{-1}_{k,k'}(t,t') = \frac{1}{2}\left[\,\partial_t^2 + \omega_k^2\,\right]\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\,\delta_{kk'}\,\delta(t,t')\,, \tag{17.83}$$

and

$$D^{-1}_{k,k'}(t,t') = \begin{pmatrix}i\partial_t - \epsilon_k & 0\\ 0 & -i\partial_t - \epsilon_k\end{pmatrix}\,\delta_{kk'}\,\delta(t,t')\,. \tag{17.84}$$

Here $\phi_k(t)$ are Bose fields and $\chi_k(t)$ Fermi fields. We can also define the four-component supervector:

$$\Phi_k(t) = \begin{pmatrix}\phi_k(t)\\ \chi_k(t)\end{pmatrix}\,,\qquad \Phi_k^\dagger(t) = \begin{pmatrix}\phi_k^*(t), & \chi_k^*(t)\end{pmatrix}\,, \tag{17.85}$$

and the inverse Green function matrix:

$$\mathcal{G}^{-1}_{k,k'}(t,t') = \begin{pmatrix}G^{-1}_{k,k'}(t,t') & 0\\ 0 & D^{-1}_{k,k'}(t,t')\end{pmatrix}\,. \tag{17.86}$$

Then the action can be written in a very compact way as:

$$S[\Phi] = -\frac{1}{2}\int \mathrm{d}t \int \mathrm{d}t' \frac{1}{N^2}\sum_{k,k'=-[N/2]}^{[N/2]}\Big\{\,\Phi_k^\dagger(t)\,\mathcal{G}^{-1}_{k,k'}(t,t')\,\Phi_{k'}(t')$$
$$+ V_{k,k'}\,\tilde{\phi}_{k-k'}(t)\,\tilde{\psi}_k^*(t)\,\tilde{\psi}_{k'}(t)\,\Big\}\,. \tag{17.87}$$

Since the action is *quadratic* in the $\tilde{\psi}_k(t)$-variables, an approximation in which we integrate away these variables in favor of an effective action in terms of the $\phi_k(t)$ variables suggests itself. We invesitgate this approximation in a following section.

17.7 Quantization

For the electron system, the Grassmann canonical variables ψ_n, ψ_n^* become quantum operators, obeying anticommutation relations. These are given by:

$$\{\,\psi_n, \psi_{n'}^\dagger\,\} = \delta_{n,n'}\,, \qquad \{\,\psi_n, \psi_{n'}\,\} = \{\,\psi_n^\dagger, \psi_{n'}^\dagger\,\} = 0\,. \tag{17.88}$$

The Fourier transformed operators obey:

$$\{\,\tilde{\psi}_k, \tilde{\psi}_{k'}^\dagger\,\} = \delta_{k,k'}\,, \tag{17.89}$$

with all other operators anticommuting. The electron number operator is given by:

$$N_{\rm e} = \sum_{n=0}^{N-1} \phi_n^\dagger(t)\,\phi_n(t) = \frac{1}{N}\sum_{k=-[N/2]}^{[N/2]} \tilde{\psi}_k^\dagger(t)\,\tilde{\psi}_k(t)\,. \tag{17.90}$$

For the lattice motion, ϕ_n and $\pi_n = M\,\dot{\phi}_n$ are real conjugate variables which become Hermitian operators in quantum mechanics and obey the commutation relations:

$$[\,\phi_n(t), \pi_{n'}(t)\,] = i\hbar\,\delta_{n,n'}\,. \tag{17.91}$$

The Fourier transformed non-Hermitian operators $a_k(t)$ and $a_{k'}^\dagger(t)$ obey:

$$[\,a_k(t), a_{k'}^\dagger(t)\,] = \delta_{k,k'}\,, \tag{17.92}$$

with all other phonon operators commuting.

17.8 Block wave functions

An electron moving in a periodic one-dimensional lattice experiences a periodic potential. We have written this potential in the form:

$$V(x) = \sum_{n=0}^{N-1} V(x - an)\,. \tag{17.93}$$

Our approximation in the last section has been to expand the electron wave function in a basis set of wave functions localized at the the N sites. We do not need to do this. Instead, we can expand the wave function directly in terms of solutions of the electron in the complete periodic potential. Solutions of the Schrödinger equation in a periodic potential are called "Block wave functions," and were studied by Felix Block in the early 1930's in his investigation of the conduction of electricity in metals.

17.8.1 A one-dimensional periodic potential

Block wave functions are solutions of Schrödinger's equation

$$\left\{\,\frac{\hbar^2}{2m}\,\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\,\right\}\psi(x) = E\,\psi(x)\,, \tag{17.94}$$

for the case of a periodic potential, $V(x+a) = V(x)$, with period a. The solution can be stated as a theorem, called "Floquet's theorem:"

Theorem 37 (Floquet's theorem)**.** *The solution of Schrödinger's Eq. (17.94) for a periodic potential of period a can be expressed as:*

$$\psi_K(x) = e^{iKx}\,u_K(x)\,, \tag{17.95}$$

where $u_K(x+a) = u_K(x)$ is a periodic function with period a.

Proof. If we put $x \to x + a$ in Eq. (17.94), we see that $\psi(x + a)$ satisfies:

$$\left\{ \frac{\hbar^2}{2m} \frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x) \right\} \psi(x + a) = E\, \psi(x + a)\,. \tag{17.96}$$

So this means that $\psi(x+a)$ is the solution of the same equation as before and therefore, since the probability of finding the electron somewhere must be the same, $\psi(x + a)$ can differ from $\psi(x)$ by only a phase:

$$\psi(x + a) = e^{iKa}\, \psi(x)\,, \tag{17.97}$$

where we have chosen the phase to be Ka. The solution of (17.97) can be expressed in the form (17.95), which completes the proof. □

17.8.2 A lattice of delta-functions

We find here solutions of the one-dimensional Schrödinger's time-independent equation (17.94) for a particle in a periodic delta-function potential of the form:

$$V(x) = -\lambda \sum_n \delta(x - an)\,, \tag{17.98}$$

where a is the lattice spacing. By Floquet's theorem, the solution of Schrödinger's equation is given by:

$$\psi_K(x) = e^{iKx}\, u_K(x)\,, \qquad \text{where} \qquad u_K(x + a) = u_K(x)\,. \tag{17.99}$$

Periodic boundary conditions at $x = 0$ and $x = L = Na$ give:

$$K_n = \frac{2\pi n}{L} = \frac{2\pi n}{aN}\,, \qquad \text{for } n = 0, 1, 2, \ldots, N - 1. \tag{17.100}$$

So we find that:

$$\begin{aligned} \psi(x + a) &= e^{iK(x+a)}\, u_K(x) = e^{iKa}\, \psi(x)\,, \\ \psi(x - a) &= e^{iK(x-a)}\, u_K(x) = e^{-iKa}\, \psi(x)\,, \end{aligned} \tag{17.101}$$

so that we only need to solve for $\psi(x)$ in the range $0 < x < a$, and then use the second of Eq. (17.101) to find $\psi(x)$ in the range $-a < x < 0$. That is, for positive energies, $E = \hbar^2 k^2/(2m)$, the solutions are given by:

$$\psi(x) = \begin{cases} A\, \cos(kx) + B\, \sin(kx)\,, & \text{for } 0 < x < a, \\ e^{-iKa} \left\{ A\, \cos(k(x + a)) + B\, \sin(k(x + a)) \right\}, & \text{for } -a < x < 0, \end{cases} \tag{17.102}$$

whereas for negative energies, $E = -\hbar^2 \kappa^2/(2m)$, we find:

$$\psi(x) = \begin{cases} A\, \cosh(\kappa x) + B\, \sinh(\kappa x)\,, & \text{for } 0 < x < a, \\ e^{-iKa} \left\{ A\, \cosh(\kappa(x + a)) + B\, \sinh(\kappa(x + a)) \right\}, & \text{for } -a < x < 0. \end{cases} \tag{17.103}$$

The boundary conditions require that the wave function be continuous at $x = 0$ and that the derivative of the wave function be discontinuous, with a jump given by:

$$\frac{\mathrm{d}\psi(+\epsilon)}{\mathrm{d}x} - \frac{\mathrm{d}\psi(-\epsilon)}{\mathrm{d}x} = -\frac{2m\lambda}{\hbar^2}\, \psi(0)\,. \tag{17.104}$$

This gives the two equations:

$$\begin{aligned} A &= e^{-iKa} \left\{ A\, \cos(ka) + B\, \sin(ka) \right\}, \\ B - e^{-iKa} \left\{ -A\, \sin(ka) + B\, \cos(ka) \right\} &= -2\beta\, A/(ka)\,, \end{aligned}$$

for positive energies, and

$$A = e^{-iKa}\left\{ A\,\cosh(\kappa a) + B\,\sinh(\kappa a)\right\},$$
$$B - e^{-iKa}\left\{ A\,\sinh(\kappa a) + B\,\cosh(\kappa a)\right\} = -2\beta\,A/(\kappa a),$$

for negative energies, where $\beta = m\lambda a/\hbar^2$. The boundary conditions give two equations for A and B:

$$\begin{aligned}\left[e^{-iKa}\,\cos(ka) - 1\right]A + e^{-iKa}\,\sin(ka)\,B &= 0\,,\\ \left[e^{-iKa}\,\sin(ka) + 2\beta/(ka)\right]A + \left[1 - e^{-iKa}\,\cos(ka)\right]B &= 0\,,\end{aligned} \tag{17.105}$$

for positive energies, and

$$\begin{aligned}\left[e^{-iKa}\,\cosh(\kappa a) - 1\right]A + e^{-iKa}\,\sinh(\kappa a)\,B &= 0\,,\\ \left[-e^{-iKa}\,\sinh(\kappa a) + 2\beta/(\kappa a)\right]A + \left[1 - e^{-iKa}\,\cosh(\kappa a)\right]B &= 0\,,\end{aligned} \tag{17.106}$$

for negative energies. Non-trivial solutions exist when the determinant of these equations vanish. This give the eigenvalue equations:

$$\cos(K_n a) = \cos(ka) - \beta\,\sin(ka)/(ka)\,. \tag{17.107}$$

for positive energies, and

$$\cos(K_n a) = \cosh(\kappa a) - \beta\,\sinh(\kappa a)/(\kappa a)\,. \tag{17.108}$$

for negative energies. Numerical solutions of these equations are shown in Figs. 17.9, 17.10, and 17.11 for $\beta = 0.5$, 1.0, and 1.5. Here values of ka/π are plotted on the positive x-axis and values of $\kappa a/\pi$ on the negative x-axis. Solutions exists for values of k and κ for values of the curve which lie between -1 and $+1$. The energy levels, shown in Fig. 17.12 for $\beta = 1.5$, therefore shows a band structure, with energy gaps between the bands.

17.8.3 Numerical methods

In this section, we show how to numerically solve Schrödinger's equation for a particle in a one-dimensional periodic potential. The solutions are called Block wave functions. The eigenvalues for this problem give a band structure with energy gaps.

We follow the work of Reitz [**?**]. We consider a one dimensional lattice with lattice spacing a, such that the potential is periodic in a, $V(x+a) = V(x)$, and a region of space between zero and L, which is a multiple of the lattice spacing a, that is: $0 \le x \le L = Ma$. Block wave functions $\psi_k(x)$ are solutions of Schrödinger's time-independent energy eigenvalue equation for this periodic potential:

$$\left\{ -\frac{1}{2m}\,\frac{\partial^2}{\partial x^2} + V(x)\right\}\psi_k(x) = \epsilon_k\,\psi_k(x)\,. \tag{17.109}$$

By Floquet's theorem, the solutions of this problem is of the form:

$$\psi_k(x) = e^{ikx}\,u_k(x)\,, \tag{17.110}$$

with $u_k(x+a) = u_k(x)$ periodic. $u_k(x)$ satisfies the differential equation:

$$\left\{ \frac{1}{2m}\left[\frac{1}{i}\,\frac{\partial}{\partial x} + k\right]^2 + V(x)\right\}u_k(x) = \epsilon_k\,u_k(x)\,. \tag{17.111}$$

We also demand that the full region between zero and L be periodic. This means that we require $\psi_k(L) = \psi_k(0)$, or

$$e^{ikL}\,u_k(Ma) = u_k(0)\,. \tag{17.112}$$

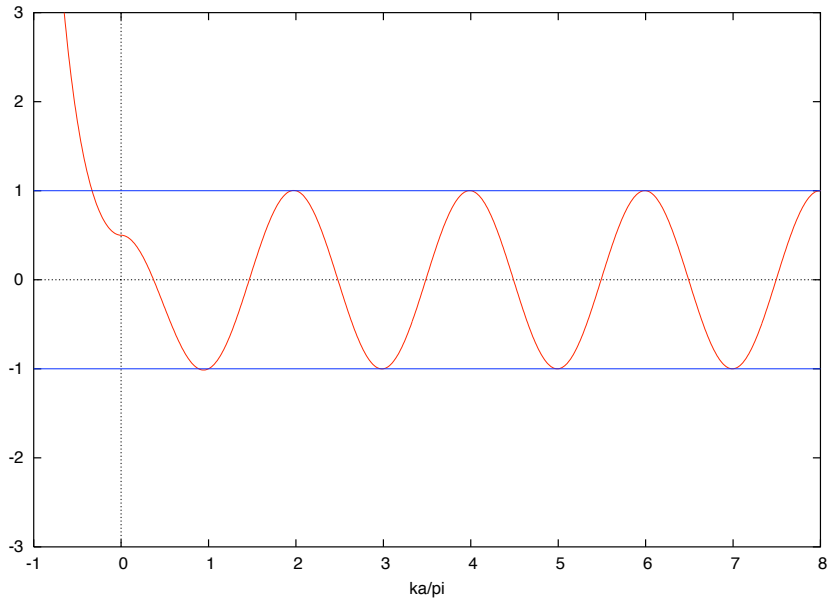

Figure 17.9: Plot of the right-hand side of Eqs. (17.107) and (17.108), for $\beta = 0.5$.

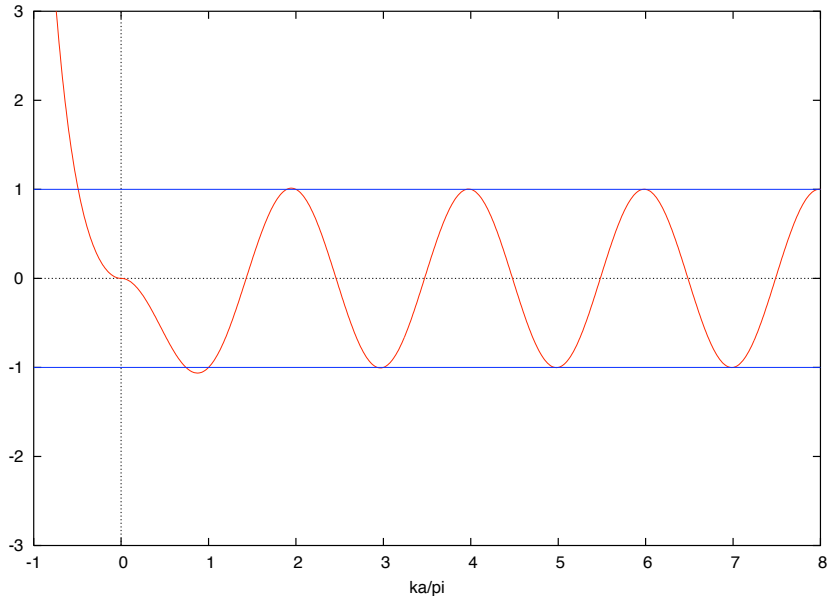

Figure 17.10: Plot of the right-hand side of Eqs. (17.107) and (17.108), for $\beta = 1.0$.

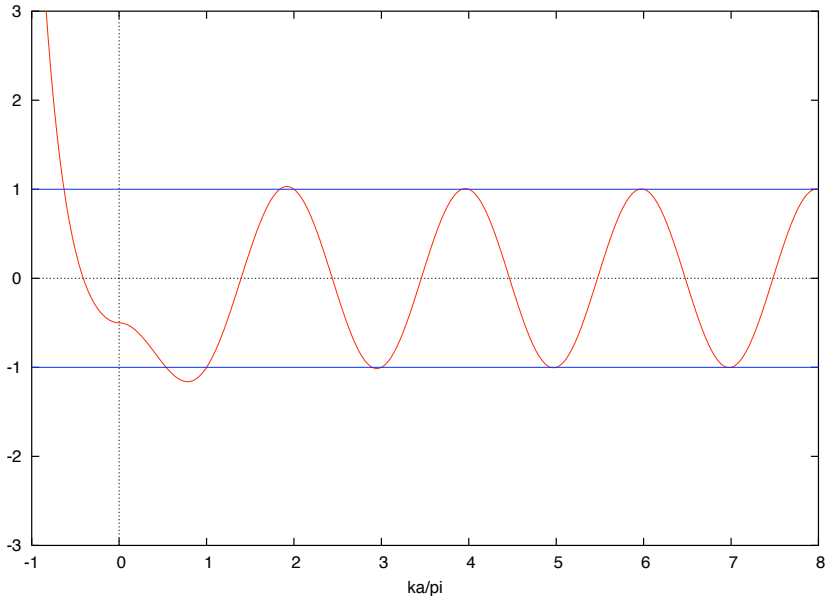

Figure 17.11: Plot of the right-hand side of Eqs. (17.107) and (17.108), for $\beta = 1.5$.

But since $u_k(Ma) = u_k(0)$, we must have:

$$k = \frac{2\pi n}{L} = \frac{2\pi n}{aM}\,, \qquad \text{for } n = 0, 1, 2, \ldots, M-1. \tag{17.113}$$

The boundary conditions on $u_k(x)$ are:

$$u_k(a) = u_k(0)\,, \qquad u_k'(a) = u_k'(0)\,. \tag{17.114}$$

Translating these boundary conditions on $u_k(x)$ to boundary conditions on $\psi_k(x)$, we find:

$$\begin{aligned} \psi_k(a) &= \psi_k(0)\, e^{ika}\,, \\ \Big\{\, \psi_k'(a) - ik\,\psi_k(a) \,\Big\} &= \Big\{\, \psi_k'(0) - ik\,\psi_k(0) \,\Big\}\, e^{ika}\,. \end{aligned} \tag{17.115}$$

These equations can be combined to give:

$$\begin{aligned} \psi_k(a) &= \psi_k(0)\, e^{ika}\,, \\ \psi_k'(a) &= \psi_k'(0)\, e^{ika}\,. \end{aligned} \tag{17.116}$$

It is easier to solve Eq. (17.109) for $\psi_k(x)$ that Eq. (17.111) for $u_k(x)$. The energy ϵ_k is then determined by the solution of Eq. (17.109) in the region $0 \le x \le a$, subject to boundary conditons (17.116). The energy is generally a *multivalued* function of k, so it will become useful later to label energies and wave functions by an additional label indicating the k-branch of the energy function.

Block wave functions are orthogonal over the full region $0 \le x \le L$. From (17.109), we find:

$$\begin{aligned} \psi_{k'}^*(x)\Big\{ -\frac{1}{2m}\,\frac{\partial^2}{\partial x^2} + V(x) \Big\}\,\psi_k(x) &= \epsilon_k\,\psi_{k'}^*(x)\,\psi_k(x)\,, \\ \psi_k(x)\Big\{ -\frac{1}{2m}\,\frac{\partial^2}{\partial x^2} + V(x) \Big\}\,\psi_{k'}^*(x) &= \epsilon_{k'}\,\psi_k(x)\,\psi_{k'}^*(x)\,. \end{aligned}$$

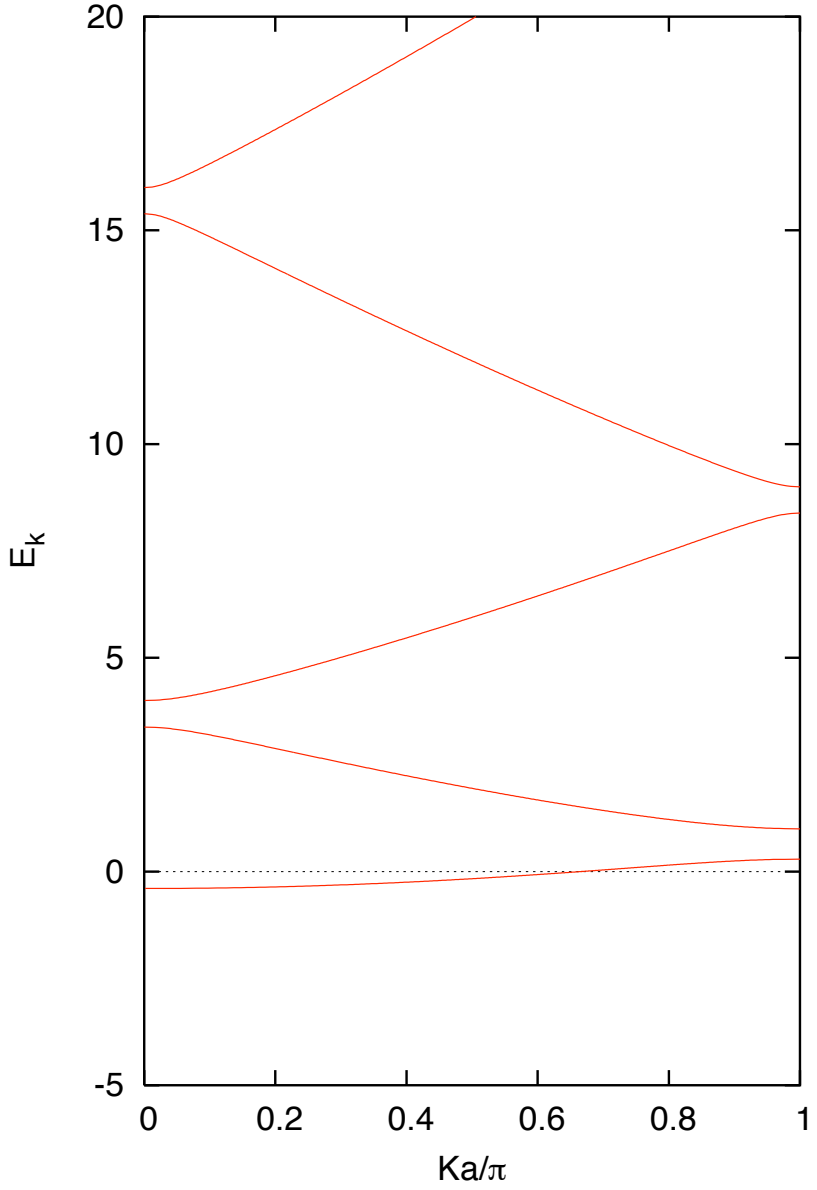

Figure 17.12: Plot of the energy, in units of $2m/\hbar^2$, as a function of Ka/π for $\beta = 1.5$.

Subtracting these two equations and integrating over x gives:

$$(\epsilon_k - \epsilon_{k'}) \int_0^L \mathrm{d}x\, \psi^*_{k'}(x)\, \psi_k(x) = \frac{1}{2m}\left[\psi^*_{k'}(x)\,\psi'_k(x) - \psi^{*\prime}_{k'}(x)\,\psi_k(x)\right]_0^L = 0\,, \tag{17.117}$$

for $\epsilon_k \neq \epsilon_{k'}$, since $\psi_k(l) = \psi_k(0)$ and $\psi'_k(L) = \psi'_k(0)$, for all k. The wave functions for the *same* values of k but on *different* k-branches of the energy function are also orthogonal. We can normalize them in the full region $0 \leq x \leq L$, and require:

$$\int_0^L \mathrm{d}x\, \psi^*_{n,b}(x)\, \psi_{n',b'}(x) = \delta_{n,n'}\,\delta_{b,b'}\,. \tag{17.118}$$

Here we have labeled the wave functions by the value of n which determines k_n and the k-branch, which we label by $b = 1, 2, \ldots$. So we can always expand the field $\psi(x,t)$ in Block wave functions:

$$\psi(x,t) = \sum_{n,b} q_{n,b}(t)\,\psi_{n,b}(x)\,, \tag{17.119}$$

so that:

$$\begin{aligned} H_0 &= \int_0^L \mathrm{d}x\, \psi^*(x,t)\left\{-\frac{1}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right\}\psi(x,t) = \sum_{n,b}\epsilon_{n,b}\,|q_{n,b}(t)|^2\,, \\ N &= \int_0^L \mathrm{d}x\,|\psi(x,t)|^2 = \sum_{n,b}|q_{n,b}(t)|^2\,. \end{aligned} \tag{17.120}$$

We can always solve Eq. (17.109) for $\psi_k(x)$ numerically. We review here how to do that. We first translate the region $0 \leq x \leq a$ into a region symmetric about the origin by setting $x \to x - a/2$. The differential equation (17.109) remains the same, but the boundary conditions (17.116) now become:

$$\begin{aligned} \psi_k(a/2)\,e^{-ika/2} &= \psi_k(-a/2)\,e^{+ika/2}\,, \\ \psi'_k(a/2)\,e^{-ika/2} &= \psi'_k(-a/2)\,e^{+ika/2}\,. \end{aligned} \tag{17.121}$$

We now write $\psi_k(x)$ as a sum of even and odd functions. We let:

$$\psi_k(x) = A_k\, f_k(x) + i\, B_k\, g_k(x)\,, \tag{17.122}$$

where $f_k(-x) = f_k(x)$ and $g_k(-x) = -g_k(x)$. Starting at the origin, we choose a value of ϵ_k and numerically integrate. using the differential equation, the even and odd functions separately out to $x = a/2$, then compute the values and first derivatives of f and g. We can take f and g to be real. The boundary conditions (17.121) become:

$$\begin{aligned} \sin(ka/2)\,f_k(a/2)\,A_k - \cos(ka/2)\,g_k(a/2)\,B_k &= 0\,, \\ \cos(ka/2)\,f'_k(a/2)\,A_k + \sin(ka/2)\,g'_k(a/2)\,B_k &= 0\,, \end{aligned} \tag{17.123}$$

which have solutions only if:

$$\frac{B_k}{A_k} = \frac{\sin(ka/2)}{\cos(ka/2)}\,\frac{f_k(a/2)}{g_k(a/2)} = -\frac{\cos(ka/2)}{\sin(ka/2)}\,\frac{f'_k(a/2)}{g'_k(a/2)}\,, \tag{17.124}$$

or

$$\tan(ka/2) = \sqrt{-\frac{g_k(a/2)\,f'_k(a/2)}{g'_k(a/2)\,f_k(a/2)}}\,. \tag{17.125}$$

Since the wave functions on the right-hand side depend on the value of ϵ_k from the differential equation, Eq. (17.124) *determines* the eigenvalues ϵ_k for a given value of k, provided k is *real*. There can be many

values of ϵ_k for a given value of k. The edges of the allowed bands are when either $f_k(a/2)$, $f'_k(a/2)$, $g_k(a/2)$, or $g'_k(a/2)$ is zero. We would expect that the $k = 0$ state in the first band (the ground state) is when $f'_k(a/2) = 0$. The edges of the band are when:

$$\frac{ka}{2} = \frac{\pi n}{M} = \frac{\pi}{2}\, m\,, \qquad \text{with } m = 0, 1, 2, \ldots\,. \tag{17.126}$$

or $n = m\, M/2$. However, since $0 \le n \le M - 1$, m is restricted to be either 0 or 1. Once a solution is found which satisfies (17.125), the Block wave function and it's derivative is given by the equations:

$$\begin{aligned} \psi_k(x) &= \mathcal{N}_k \left\{ \cos(ka/2)\, \frac{f_k(x)}{f_k(a/2)} + i\, \sin(ka/2)\, \frac{g_k(x)}{g_k(a/2)} \right\}, \\ \psi'_k(x) &= \mathcal{N}'_k \left\{ \cos(ka/2)\, \frac{g'_k(x)}{g'_k(a/2)} + i\, \sin(ka/2)\, \frac{f'_k(x)}{f'_k(a/2)} \right\}, \end{aligned} \tag{17.127}$$

where

$$\begin{aligned} \mathcal{N}_k &= \frac{f_k(a/2)}{\cos(ka/2)}\, A_k\,, \\ \mathcal{N}'_k &= -i\, \frac{f'_k(a/2)}{\sin(ka/2)}\, A_k\,. \end{aligned} \tag{17.128}$$

So

$$\frac{\mathcal{N}_k}{\mathcal{N}'_k} = \sqrt{\frac{f_k(a/2)\, g_k(a/2)}{f'_k(a/2)\, g'_k(a/2)}}\,. \tag{17.129}$$

Recall that our normalization is:

$$M \int_{-a/2}^{a/2} \mathrm{d}x\, |\psi_k(x)|^2 = 1\,. \tag{17.130}$$

So, from the first of (17.128), we can fix $\mathcal{N}_k$ by the numerical requirement:

$$\frac{1}{|\mathcal{N}_k|^2} = 2M \int_0^{a/2} \mathrm{d}x \left\{ \cos^2(ka/2) \left[\frac{f_k(x)}{f_k(a/2)} \right]^2 + \sin^2(ka/2) \left[\frac{g_k(x)}{g_k(a/2)} \right]^2 \right\}. \tag{17.131}$$

This completes the discussion of the numerical solution of the Block equation.

Example 32. Atoms can be trapped by potentials, called "electromagnetic tweezers," created by laser beams. These potentials are of the form:

$$V(x) = \frac{V_0}{2} \left[1 - \cos\left(\frac{2\pi x}{a} \right) \right]. \tag{17.132}$$

The eigenvalue equation (17.109) then becomes:

$$\left\{ -\frac{1}{2m}\, \frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{V_0}{2} \left[1 - \cos\left(\frac{2\pi x}{a} \right) \right] \right\} \psi_k(x) = \epsilon_k\, \psi_k(x)\,. \tag{17.133}$$

Changing variables, we set:

$$z = \frac{\pi x}{a}\,, \qquad s = \frac{2mV_0a^2}{\pi^2}\,, \qquad e_k = \frac{a\epsilon_k}{\pi} \sqrt{\frac{2m}{V_0}}\,. \tag{17.134}$$

then (17.133) becomes:

$$\left\{ -\frac{\mathrm{d}^2}{\mathrm{d}z^2} + \frac{s}{2} \left[1 - \cos(2z) \right] \right\} \psi_k(z) = e_k\, \sqrt{s}\, \psi_k(z)\,. \tag{17.135}$$

This is one form of the Mathieu equation [2][p. 720], the solutions of which are Mathieu functions.

References

[1] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN: The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 1992).

[2] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).

Chapter 18

Schrödinger perturbation theory

In this chapter, we discuss perturbation theory in the Schrödinger representation. However we first derive a useful theorem. The Feynman-Hellman theorem states that:

Theorem 38 (Feynman-Hellman)**.** *Let λ be any real parameter of a real Hermitian operator $H(\lambda)$ satisfying the eigenvalue equation:*

$$H(\lambda)\,|\,\psi_{n,\alpha}(\lambda)\,\rangle = E_n(\lambda)\,|\,\psi_{n,\alpha}(\lambda)\,\rangle\,, \tag{18.1}$$

where $|\,\psi_{n,\alpha}(\lambda)\,\rangle$ is any of the eigenvectors with eigenvalue $E_n(\lambda)$. Then the rate of change of the eigenvalue $E_n(\lambda)$ with λ is given by:

$$\frac{\partial E_n(\lambda)}{\partial\lambda} = \left\langle\,\psi_{n,\alpha}(\lambda)\,\right|\frac{\partial H(\lambda)}{\partial\lambda}\left|\,\psi_{n,\alpha}(\lambda)\,\right\rangle. \tag{18.2}$$

Proof. Differentiating Eq. (18.1) by λ gives:

$$\frac{\partial H(\lambda)}{\partial\lambda}\,|\,\psi_{n,\alpha}(\lambda)\,\rangle + H(\lambda)\,\frac{\partial|\,\psi_{n,\alpha}(\lambda)\,\rangle}{\partial\lambda} = \frac{\partial E_n(\lambda)}{\partial\lambda}\,|\,\psi_{n,\alpha}(\lambda)\,\rangle + E(\lambda)\,\frac{\partial|\,\psi_{n,\alpha}(\lambda)\,\rangle}{\partial\lambda}\,. \tag{18.3}$$

Multiplying through by $\langle\,\psi_{n,\alpha}(\lambda)\,|$ and using the fact that H is Hermitian, results in cancelation of the derivatives of the eigenvectors and in Eq. (18.2), which was what we wanted to prove. The result is exact. □

18.1 Time-independent perturbation theory

In this section, we derive equations for first and second order time-independent perturbation theory for operators. Suppose we wish to find an approximate solution to the problem:

$$H(\lambda)\,|\,\psi(\lambda)\,\rangle = E(\lambda)\,|\,\psi(\lambda)\,\rangle\,, \tag{18.4}$$

where $H(\lambda)$ is given by: $H(\lambda) = H_0 + \lambda V$ with λ, in some sense, "small." Then it is clearly advantagious to try to expand the eigenvector and eigenvalue in a power series in λ:

$$\begin{aligned} |\,\psi(\lambda)\,\rangle &= \sum_{\alpha=1}^{M(n)} c^{(0)}_{n,\alpha}\,|\,\psi^{(0)}_{n,\alpha}\,\rangle + \lambda\sum_{\alpha=1}^{M(n)} c^{(1)}_{n,\alpha}\,|\,\psi^{(1)}_{n,\alpha}\,\rangle + \lambda^2\sum_{\alpha=1}^{M(n)} c^{(2)}_{n,\alpha}\,|\,\psi^{(2)}_{n,\alpha}\,\rangle + \cdots\,, \\ E(\lambda) &= E^{(0)}_n + \lambda\,E^{(1)}_n + \lambda^2\,E^{(2)}_n + \cdots\,, \end{aligned} \tag{18.5}$$

where $|\,\psi^{(0)}_{n,\alpha}\,\rangle$ is a solution of:

$$H_0\,|\,\psi^{(0)}_{n,\alpha}\,\rangle = E^{(0)}_n\,|\,\psi^{(0)}_{n,\alpha}\,\rangle\,, \qquad \langle\,\psi^{(0)}_{n,\alpha}\,|\,\psi^{(0)}_{n',\alpha'}\,\rangle = \delta_{n,n'}\,\delta_{\alpha,\alpha'}\,, \tag{18.6}$$

with α labeling the $M(n)$ possible degenerate states of the unperturbed Hamiltonian H_0. The coefficients $c^{(m)}_{n,\alpha}$ are to be determined. They are normalized so that:

$$\sum_{\alpha=1}^{M(n)} |\, c^{(m)}_{n,\alpha}\,|^2 = 1\,. \tag{18.7}$$

For the case when there are no degeneracies, $M(n) = 1$ and we can set all the $c^{(m)}_{n,\alpha} = 1$. We considere here the general case for any value of $M(n)$. Substituting (18.5) into (18.4) gives:

$$\begin{aligned}\{\, H_0 + \lambda\, V\,\} \sum_{\alpha=1}^{M(n)} \{\, c^{(0)}_{n,\alpha}\,|\,\psi^{(0)}_{n,\alpha}\,\rangle + \lambda\, c^{(1)}_{n,\alpha}\,|\,\psi^{(1)}_{n,\alpha}\,\rangle + \lambda^2\, c^{(2)}_{n,\alpha}\,|\,\psi^{(2)}_{n,\alpha}\,\rangle + \cdots\,\} \\ = \{\, E^{(0)}_n + \lambda\, E^{(1)}_n + \lambda^2\, E^{(2)}_n + \cdots\,\} \sum_{\alpha=1}^{M(n)} \{\, c^{(0)}_{n,\alpha}\,|\,\psi^{(0)}_{n,\alpha}\,\rangle + \lambda\, c^{(1)}_{n,\alpha}\,|\,\psi^{(1)}_{n,\alpha}\,\rangle + \lambda^2\, c^{(2)}_{n,\alpha}\,|\,\psi^{(2)}_{n,\alpha}\,\rangle + \cdots\,\}\,.\end{aligned}$$

Equating coefficients of λ yields the equations:

$$H_0\,|\,\psi^{(0)}_{n,\alpha}\,\rangle = E^{(0)}_n\,|\,\psi^{(0)}_{n,\alpha}\,\rangle\,, \tag{18.8}$$

$$\sum_{\alpha=1}^{M(n)} \{\, c^{(1)}_{n,\alpha}\, H_0\,|\,\psi^{(1)}_{n,\alpha}\,\rangle + c^{(0)}_{n,\alpha}\, V\,|\,\psi^{(0)}_{n,\alpha}\,\rangle\,\} = \sum_{\alpha=1}^{M(n)} \{\, c^{(1)}_{n,\alpha}\, E^{(0)}_n\,|\,\psi^{(1)}_{n,\alpha}\,\rangle + c^{(0)}_{n,\alpha}\, E^{(1)}_n\,|\,\psi^{(0)}_{n,\alpha}\,\rangle\,\}\,, \tag{18.9}$$

$$\begin{aligned}\sum_{\alpha=1}^{M(n)} \{\, c^{(2)}_{n,\alpha}\, H_0\,|\,\psi^{(2)}_{n,\alpha}\,\rangle + c^{(1)}_{n,\alpha}\, V\,|\,\psi^{(1)}_{n,\alpha}\,\rangle\,\} = \sum_{\alpha=1}^{M(n)} \{\, c^{(2)}_{n,\alpha}\, E^{(0)}_n\,|\,\psi^{(2)}_{n,\alpha}\,\rangle + c^{(1)}_{n,\alpha}\, E^{(1)}_n\,|\,\psi^{(1)}_{n,\alpha}\,\rangle \\ + c^{(0)}_{n,\alpha}\, E^{(2)}_n\,|\,\psi^{(0)}_{n,\alpha}\,\rangle\,\}\,,\end{aligned} \tag{18.10}$$

$$\cdots\cdots$$

Eq. (18.8) defines the unperturbed solutions (18.6). Operating on (18.9) on the left by $\langle\,\psi^{(0)}_{n',\alpha'}\,|$ and using the Hermitian property of H_0 gives:

$$\begin{aligned}\sum_{\alpha=1}^{M(n)} \left(\, E^{(0)}_n - E^{(0)}_{n'}\,\right) \langle\,\psi^{(0)}_{n',\alpha'}\,|\,\psi^{(1)}_{n,\alpha}\,\rangle\, c^{(1)}_{n,\alpha} \\ + \sum_{\alpha=1}^{M(n)} \left\{\, E^{(1)}_n\,\langle\,\psi^{(0)}_{n',\alpha'}\,|\,\psi^{(0)}_{n,\alpha}\,\rangle - \langle\,\psi^{(0)}_{n',\alpha'}\,|\,V\,|\,\psi^{(0)}_{n,\alpha}\,\rangle\,\right\} c^{(0)}_{n,\alpha} = 0\,.\end{aligned} \tag{18.11}$$

Setting $n' = n$, the first term in (18.11) vanishes, leaving the set of equations:

$$\sum_{\alpha=1}^{M(n)} \left\{\, E^{(1)}_n\,\delta_{\alpha',\alpha} - \langle\,\psi^{(0)}_{n',\alpha'}\,|\,V\,|\,\psi^{(0)}_{n,\alpha}\,\rangle\,\right\} c^{(0)}_{n,\alpha} = 0\,. \tag{18.12}$$

The set of Eqs. (18.12) has $M(n)$ eigenvalues $E^{(1)}_n(\beta)$ and eigenvectors $c^{(0)}_{n,\alpha}(\beta)$, which we label by $\beta = 1,\ldots,M(n)$. This equation fixes the values of $c^{(0)}_{n,\alpha}(\beta)$. We define eigenvectors $|\,\phi^{(0)}_{n,\beta}\,\rangle$ of the degenerate state n which diagonalize the matrix (18.12) by:

$$|\,\phi^{(0)}_{n,\beta}\,\rangle = \sum_{\alpha=1}^{M(n)} c^{(0)}_{n,\alpha}(\beta)\,|\,\psi^{(0)}_{n,\alpha}\,\rangle\,. \tag{18.13}$$

The set of vectors $|\,\phi^{(0)}_{n\,,\beta}\,\rangle$ are also orthonormal:

$$\langle\,\phi^{(0)}_{n\,,\beta}\,|\,\phi^{(0)}_{n',\beta'}\,\rangle = \delta_{n,n'}\,\delta_{\beta,\beta'}\,. \tag{18.14}$$

Now setting $n' \neq n$ in Eq. (18.11), the second term in (18.11) vanishes, from which we find the result for the overlap:

$$\langle\,\phi^{(0)}_{n',\beta'}\,|\,\phi^{(1)}_{n\,,\beta}\,\rangle = \frac{\langle\,\phi^{(0)}_{n',\beta'}\,|\,V\,|\,\phi^{(0)}_{n\,,\beta}\,\rangle}{E^{(0)}_n - E^{(0)}_{n'}}\,, \qquad n' \neq n\,, \tag{18.15}$$

where we have defined:

$$|\,\phi^{(1)}_{n,\beta}\,\rangle = \sum_{\alpha=1}^{M(n)} c^{(1)}_{n,\alpha}(\beta)\,|\,\psi^{(1)}_{n,\alpha}\,\rangle\,, \tag{18.16}$$

and have multiplied by $c^{(0)}_{n',\alpha'}(\beta')$ and summed over α'. So, to first order, the eigenvectors and eigenvalues are given by:

$$\begin{aligned} |\,\psi_n(\beta)\,\rangle &= |\,\phi^{(0)}_{n\,,\beta}\,\rangle + \sum_{\substack{\beta'=1\\ n'\neq n}}^{M(n)} \frac{|\,\phi^{(0)}_{n',\beta'}\,\rangle\,\langle\,\phi^{(0)}_{n',\beta'}\,|\,\lambda\,V\,|\,\phi^{(0)}_{n\,,\beta}\,\rangle}{E^{(0)}_n - E^{(0)}_{n'}}\,, \\ E_n(\beta) &= E^{(0)}_n + \lambda\,E^{(1)}_n(\beta)\,. \end{aligned} \tag{18.17}$$

This eigenvector is normalized to order λ^2.

Remark 26. The question of whether the perturbative power series in λ for the eigenvalues and eigenvectors converge cannot be answered in general. In general, the perturbation series does not lead to a normalized eigenvector. In order to correct for this, and perhaps for other problems with perturbation theory, one often tries to resum parts of the perturbation series. This has produced a proliferation of resumed "non-perturbative" approximations to perturbation theory. One of these, for example is the XXX approximation.

18.2 Time-dependent perturbation theory

In this section, we derive equations for time-dependent perturbation theory in the Schrödinger representation.

References

Chapter 19

Variational methods

19.1 Introduction

Use of variational approximations in quantum mechanics have a long and interesting history. The time-*independent* method was first applied by Lord Rayleigh in 1873 to the computation of the vibration frequencies of mechanical systems[**?**]. In this method, a normalized "trial" wave function for the ground state is taken to be a function of a number of arbitrary parameters. These parameters are varied until a minimum is found. With some ingenuity, trial wave functions with thousands of parameters have been used successfully in atomic physics for the ground states of atoms and molecules.

The time-*dependent* version of the variational approximation can be traced to an obscure appendix in the 1930 Russian edition of the "Principles of Wave Mechanics," by Dirac.[1] In this version of the variational approximation, the wave function is taken to be a function of a number of time-dependent parameters. Variation of the action, as defined by Dirac, leads to a classical set of Hamiltonian equations of motion for the parameters. These classical equations are then solved as a function of time to provide an approximation to the evolution of the wave function.

19.2 Time dependent variations

Dirac pointed out that unrestricted variation of the action:

$$S[\psi,\psi^*] = \int_0^T \mathrm{d}t \left\{ \frac{i\hbar}{2} \left[\langle\, \psi^\dagger(t) \,|\, \partial_t\, \psi(t) \,\rangle - \langle\, \partial_t\, \psi^\dagger(t) \,|\, \psi(t) \,\rangle \right] - \langle\, \psi(t) \,|\, H \,|\, \psi(t) \,\rangle \right\}, \tag{19.1}$$

with no variation at the end points $|\,\delta\psi(0)\,\rangle = |\,\delta\psi(T)\,\rangle = 0$, leads to Schrödinger's equation and its adjoint:

$$H\,|\,\psi(t)\,\rangle = i\hbar\,\partial_t |\,\psi(t)\,\rangle\,, \qquad \langle\,\psi(t)\,|\,H^\dagger = -i\hbar\,\partial_t \langle\,\psi(t)\,|\,. \tag{19.2}$$

We assume here that H is hermitian, and independent explicitly of time. Then solutions of (19.2) obey a probability conservation equation:

$$\partial_t\,\langle\,\psi(t)\,|\,\psi(t)\,\rangle = 0\,. \tag{19.3}$$

We consider in this chapter a variational approximation to the exact time-dependent wave function of the form:

$$|\,\psi(t)\,\rangle = |\,\psi(\mathcal{N}(t),\theta(t),y(t))\,\rangle \equiv \mathcal{N}(t) e^{i\theta(t)}\,|\,\psi(y(t))\,\rangle\,, \tag{19.4}$$

[1] P. A. M. Dirac, Appendix to the Russian edition of *The Principles of Wave Mechanics*, as cited by Ia. I. Frenkel, *Wave Mechanics, Advanced General Theory* (Clarendon Press, Oxford, 1934), pp. 253, 436. The reference often quoted, P. A. M. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930), does not appear to contain this equation.

where $y(t) = [\, y^1(t), y^2(t), \ldots, y^{2n}(t)\,]$ is a set of $2n$ parameters which depend on time. We have selected out two of these parameters, the normalization $\mathcal{N}(t)$ and overall phase $\theta(t)$ to treat specially. The variational approximation consists of requiring that these parameters are chosen so as to minimize the action of Eq. (19.1), subject to the constraint given by Eq. (19.3). So with the choice (19.4), the action of Eq. (19.1) becomes:

$$\begin{aligned} S[\mathcal{N},\theta,y] &= \int \mathrm{d}t\, \left\{\, -\hbar\,\dot\theta\,\mathcal{N}^2 + L(\mathcal{N},y;\dot y)\,\right\}, \\ L(\mathcal{N},y;\dot y) &= \pi_i(\mathcal{N},y)\,\dot y^i - H(\mathcal{N},y)\,, \end{aligned} \tag{19.5}$$

where

$$\begin{aligned} \pi_i(\mathcal{N},y) &= \frac{i\hbar}{2}\left\{\, \langle\, \psi(\mathcal{N},y)\,|\,\partial_i\,\psi(\mathcal{N},y)\,\rangle - \langle\, \partial_i\,\psi(\mathcal{N},y)\,|\,\psi(\mathcal{N},y)\,\rangle\,\right\}, \\ H(\mathcal{N},y) &= \langle\, \psi(\mathcal{N},y)\,|\,H\,|\,\psi(\mathcal{N},y)\,\rangle\,. \end{aligned} \tag{19.6}$$

Here we have defined:[2] $\partial_i \equiv \partial/\partial y^i$. Since the integrand of the action in Eq. (19.5) is independent of θ, Lagrange's equation for θ gives a conservation equation for the normalization:

$$\frac{\mathrm{d}\mathcal{N}^2}{\mathrm{d}t} = 0\,. \tag{19.7}$$

The integrand of the action in Eq. (19.5) is also independent of $\dot{\mathcal{N}}$, and since both $\pi_i(\mathcal{N},y)$ and $H(\mathcal{N},y)$ are proportional to $\mathcal{N}^2$, Lagrange's equation for $\mathcal{N}$ gives:

$$\hbar\,\dot\theta\,\mathcal{N}^2 = L(\mathcal{N},y) = \pi_i(\mathcal{N},y)\,\dot y^i - H(\mathcal{N},y)\,. \tag{19.8}$$

Now setting $\mathcal{N}^2 = 1$ makes $\mathcal{N} = \mathcal{N}(y)$ a function of all the y parameters, and we find:

$$\dot\theta = L(y,\dot y)/\hbar = \left\{\, \pi_i(y)\,\dot y^i - H(y)\,\right\}/\hbar\,, \tag{19.9}$$

which has the solution:

$$\theta(t) = \int_0^t \mathrm{d}t\, L(y,\dot y)/\hbar\,. \tag{19.10}$$

$\theta(t)$ can only be found *after* the equations of motion are solved for $y^i(t)$. Lagrange's equations for the y-variables are now given by:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot y^i}\right) - \frac{\partial L}{\partial y^i} = 0\,, \tag{19.11}$$

where

$$L(y,\dot y) = \pi_i(y)\,\dot y^i - H(y)\,. \tag{19.12}$$

and both $\mathcal{N}$ and θ have been eliminated in favor of y. The trial wave function is now of the form given in Eq. (19.4) with $\mathcal{N} = 1$, $\theta(t)$ given by Eq. (19.10), and normalized so that:

$$\langle\, \psi(y)\,|\,\psi(y)\,\rangle = 1\,. \tag{19.13}$$

From (19.12), the equations of motion for y are given by:

$$f_{ij}(y)\,\dot y^j = \partial_i\,H(y)\,, \qquad \text{where} \qquad f_{ij}(y) = \partial_i\,\pi_j(y) - \partial_j\,\pi_i(y)\,, \tag{19.14}$$

which can be solved if the co-variant matrix $f_{ij}(y)$ is non-singular. We define the inverse as the contra-variant matrix with upper indices:

$$f_{ij}(y)\,f^{jk}(y) = f^{ij}(y)\,f_{jk}(y) = \delta_i^j\,, \tag{19.15}$$

[2] Because of the symplectic nature of the equations of motion for y, it is natural, but *not necessary*, to use contra- and co-variant indices here.

in which case, the equations of motion can be put in the form:

$$\dot{y}^i = f^{ij}(y)\,\partial_j\,H(y) \equiv \partial^i\,H(y)\,. \tag{19.16}$$

Here, we have defined contra-variant derivatives by:

$$\partial^i \equiv f^{ij}(y)\,\partial_j\,. \tag{19.17}$$

Total energy is conserved:

$$\frac{\mathrm{d}H(y)}{\mathrm{d}t} = \dot{y}^i\,(\partial_i\,H(y)) = f^{ij}(y)\,(\partial_i\,H(y))\,(\partial_j\,H(y)) \equiv 0\,, \tag{19.18}$$

since $f^{ij}(y)$ is antisymmetric.

Definition 30. If $A(y)$ and $B(y)$ are functions of y we define Poisson brackets by:[3]

$$\{\,A(y),B(y)\,\} = (\partial_i A(y))\,f^{ij}(y)\,(\partial_j B(y)) = (\partial_i A(y))\,(\partial^i B(y))\,. \tag{19.19}$$

Note that for example $\{\,y^i,y^j\,\} = f^{ij}(y)$. However, Poisson brackets must also obey Jacobi's identity. This is proved in the following theorem.

Theorem 39 (Jacobi's identity)**.** *Poisson brackets, defined by Eq.* (19.19)*, satisfy Jacobi's identity:*

$$\{\,A,\{\,B,C\,\}\,\} + \{\,B,\{\,C,A\,\}\,\} + \{\,C,\{\,A,B\,\}\,\} = 0\,. \tag{19.20}$$

Proof. We start by noting that, after some algebra:

$$\begin{aligned}\{\,A,\{\,B,C\,\}\,\} + \{\,B,\{\,C,A\,\}\,\} + \{\,C,\{\,A,B\,\}\,\}&\\ = (\,\partial_i A\,)\,(\,\partial_j B\,)\,(\,\partial_k C\,)\,\big\{\,f^{il}(\,\partial_l f^{jk}\,) + f^{jl}(\,\partial_l f^{ki}\,) + f^{kl}(\,\partial_l f^{ij}\,)\,\big\}&\\ = (\,\partial_i A\,)\,(\,\partial_j B\,)\,(\,\partial_k C\,)\,\big\{\,(\,\partial^i f^{jk}\,) + (\,\partial^j f^{ki}\,) + (\,\partial^k f^{ij}\,)\,\big\}\,.&\end{aligned} \tag{19.21}$$

But now we note that since $f^{jk}f_{kl} = \delta^j_l$, differentiating this expression with respect to y^i, we find:

$$(\,\partial_i f^{jk}\,)\,f_{kl} + f^{jk}\,(\,\partial_i f_{kl}\,) = 0\,. \tag{19.22}$$

Inverting this expression, and interchanging indices, we find:

$$\begin{aligned}(\,\partial_i f^{jk}\,) &= -f^{jj'}\,(\,\partial_i f^{j'k'}\,)\,f^{k'k} = f^{jj'}\,f^{kk'}\,(\,\partial_i f^{j'k'}\,)\,,\\ (\,\partial^i f^{jk}\,) &= f^{ii'}\,(\,\partial_{i'} f^{jk}\,) = f^{ii'}\,f^{jj'}\,f^{kk'}\,(\,\partial_i f^{j'k'}\,)\,.\end{aligned}$$

Using this expression in the last line of Eq. (19.21), we find:

$$\begin{aligned}\{\,A,\{\,B,C\,\}\,\} + \{\,B,\{\,C,A\,\}\,\} + \{\,C,\{\,A,B\,\}\,\}&\\ = (\,\partial^i A\,)\,(\,\partial^j B\,)\,(\,\partial^k C\,)\,\big\{\,(\,\partial_i f_{jk}\,) + (\,\partial_j f_{ki}\,) + (\,\partial_k f_{ij}\,)\,\big\}\,.&\end{aligned} \tag{19.23}$$

But since

$$f_{ij}(y) = \partial_i\,\pi_j(y) - \partial_j\,\pi_i(y)\,,$$

satisfies Bianchi's identity:

$$\partial_i\,f_{jk}(y) + \partial_j\,f_{ki}(y) + \partial_k\,f_{ij}(y) = 0\,, \tag{19.24}$$

Jacobi's identity also holds for our definition of the classical Poisson brackets. This completes the proof that the set of $2n$ classical parameters are symplectic variables. □

[3] Here, we follow Das [**?**].

Remark 27. As one might expect from our use of contra- and co-variant concepts in this section, that it is possible to develop a non-metric geometry based on the symplectic structure of our equations. This geometry is throughly discussed in the book by Kramer and Saraceno [1], and we reproduce some of their results in Appendix **??**. The derivation of our variational equations are somewhat simpler using geometric concepts, but a geometric descriptions is not necessary to apply the method to any problem of interest. For that, a good guess as to the structure of the wave function is the most important question to answer. We turn next to some examples.

Remark 28. Thus we have shown, in a quite general way, that Dirac's quantum action is an extreemum for arbitrary time-dependent variational parameters of the trial state vector, when these parameters satisfy a classical symplectic (Hamiltonian) system of equations.

19.3 The initial value problem

If the initial value of the state vector is specified, the only requirement of our variational approximation is that it must match the initial wave function. That is, at $t = 0$,

$$|\,\psi(\theta(0), y(0))\,\rangle = |\,\psi(0)\,\rangle\,, \tag{19.25}$$

so that the parameterized form of the trial vector at $t = 0$ must be made to agree with the desired initial vector. This places a minor restriction on the allowed parameterization.

The solutions for $y^i(t)$, $i = 1, \ldots, 2n$ then evolve according to (19.16), with $y^i(0)$ given by initial values. For *most* of these initial values and Hamiltonians, the evolution will eventually lead to chaotic orbits. This generally signals a failure of the variational approximation so that the wave function cannot be trusted for times beyond the chaotic breakdown.[4]

19.4 The eigenvalue problem

The simplest choice of parameters y^i are those that are independent of time. For this case, variation of the action leads to equations for the variational parameters which satisfy:

$$\partial_i H(y) = 0\,, \qquad H(y) = \langle\,\psi(y)\,|\,H\,|\,\psi(y)\,\rangle\,, \tag{19.26}$$

subject to the constraint: $\langle\,\psi(y)\,|\,\psi(y)\,\rangle = 1$. This is just the time-independent variational equation, and leads to a bound on the ground state of the system, since if we expand $|\,\psi(y)\,\rangle$ in terms of the exact eigenvectors of the system,

$$|\,\psi(y)\,\rangle = \sum_n c_n\,|\,\psi_n\,\rangle\,, \tag{19.27}$$

we find that

$$H(y) = \sum_n E_n \geq E_0\,. \tag{19.28}$$

So condition (19.26) gives an upper bound on the ground state. Bounds on eigenvalues of the energy for states other than the ground state rely on constructing variational trial wave functions which are orthogonal, or approximately orthogonal, to the variational ground state. Sometimes symmetries, such as parity, can be used to find such states.

But we do not have to restruct ourselves to time-independent variational parameters to find eigenvalues of the energy. In general the solutions for $y^i(t)$ of the classical equations of motion, given by (19.16), for arbitrary initial conditions lead to chaotic orbits. However for a particular choice of initial conditions, there could be regions of phase space where periodic closed orbits exists. These stable orbits are related to eigenvalues of the quantum Hamiltonian.

[4] See ref. [**?**, **?**] for an example of this "quantum chaos."

So we suppose that it is possible to find initial conditions for the classical equations of motion which lead to periodic orbits with period T such that $y^i(T) = y^i(0)$, for all $i = 1, \ldots, 2n$. Our problem is to find these orbits. Now the exact time-dependent wave function for the n^{th} eigenvector is given by:

$$|\,\psi_n(t)\,\rangle = e^{-iE_n t/\hbar}\,|\,\psi_n\,\rangle\,, \tag{19.29}$$

and is periodic with period $T_n = 2\pi\hbar/E_n$. That is $|\,\psi_n(T_n)\,\rangle = |\,\psi_n(0)\,\rangle$. So we require that our variational wave function is also periodic with period T_n:

$$e^{i\theta(T_n)}\,|\,\psi(y(T_n))\,\rangle = e^{i\theta(0)}\,|\,\psi(y(0))\,\rangle\,. \tag{19.30}$$

But $\theta(0) = 0$, and we assume that we have found periodic orbits such that $y^i(T_n) = y^i(0)$. Then the periodic requirement on the variational wave function states that

$$\theta(T_n) = 2\pi n'\,, \tag{19.31}$$

where n' is an integer. However

$$\theta(T_n) = \int_0^{T_n} L(t)\,\frac{\mathrm{d}t}{\hbar} = \int_0^{T_n} \pi_i(t)\,\dot{y}^i(t)\,\frac{\mathrm{d}t}{\hbar} - \frac{E_n T_n}{\hbar} = \int_0^{T_n} \pi_i(t)\,\dot{y}^i(t)\,\frac{\mathrm{d}t}{\hbar} - 2\pi\,. \tag{19.32}$$

Here we have replaced the conserved energy $H(y)$ by the exact energy E_n. So we find

$$I(T_n) = \int_0^{T_n} \pi_i(t)\,\dot{y}^i(t)\,\frac{\mathrm{d}t}{2\pi\hbar} = n \tag{19.33}$$

where $n = n' + 1 \geq 0$ is a positive integer. Only certain closed orbits have integral actions, and it is these that represent approximate time-dependent variational wave functions to the eigenstates of the system. Note that these wave functions depend on time, but are periodic.

The phase space requirement (19.33) is similar to the Bohr-Sommerfeld quantization rule, however there is an essential difference: the variational "quantization rule" (19.33) applies to the action of a *classical* Hamiltonian, derived from the variational parameterization of the trial state vector, not a classical version of the *quantum* Hamiltonian, as in the usual Bohr-Sommerfeld quantization.

One way to find closed orbits with integral action is to vary the initial conditions until the action is integral, as was done by Pattanayak and Schieve[**?**].

We will study some examples of the application of these formulas in the next section.

19.5 Examples

We now turn to several example of the use of the variational approximation.

19.5.1 The harmonic oscillator

As a first example, we study the one-dimensional harmonic oscillator with the Hamiltonian:

$$H(x,p) = \frac{p^2}{2m} + \frac{1}{2}\,m\omega^2\,x^2\,, \tag{19.34}$$

and consider a trial wave function of the form:

$$\psi(x;\Gamma(t),\Sigma(t)) = \frac{1}{[\,2\pi\Gamma(t)\,]^{1/4}}\,\exp\Big\{\,i\theta(t) - x^2\,\Big[\,\frac{1}{4\,\Gamma(t)} - i\,\Sigma(t)\,\Big]\,\Big\}\,, \tag{19.35}$$

which depend on the two parameters $y^1(t) = \Gamma(t)$ and $y^2(t) = \Sigma(t)$. $\theta(t)$ is the phase parameter, defined by Eq. (19.10). The wave function is normalized, as required by the theory:

$$\int_{-\infty}^{+\infty} \mathrm{d}x \, |\, \psi(x; \Gamma(t), \Sigma(t)) \,|^2 = 1 \,. \tag{19.36}$$

After some algebra and calculus, we find the results:

$$\pi_\Gamma = 0 \,, \qquad \pi_\Sigma = -\hbar\,\Gamma \,,$$
$$H(\Gamma, \Sigma) = \frac{\hbar^2}{2m} \left\{ \frac{1}{4\Gamma} + 4\Gamma\,\Sigma^2 \right\} + \frac{1}{2}\, m\omega^2\,\Gamma \,,$$

and so

$$\begin{aligned} L(\Gamma, \Sigma; \dot{\Sigma}) &= \pi_\Gamma \dot{\Gamma} + \pi_\Sigma \dot{\Sigma} - H(\Gamma, \Sigma) \,, \\ &= -\hbar\,\Gamma\,\dot{\Sigma} - \frac{\hbar^2}{2m} \left(\frac{1}{4\Gamma} + 4\,\Gamma\,\Sigma^2 \right) - \frac{1}{2}\, m\omega^2\,\Gamma \,. \end{aligned} \tag{19.37}$$

The equations of motion are:

$$\begin{aligned} \hbar\,\dot{\Gamma} &= \frac{4\hbar^2}{m}\,\Gamma\,\Sigma \,, \\ \hbar\,\dot{\Sigma} &= \frac{\hbar^2}{2m} \left(\frac{1}{4\Gamma^2} - 4\,\Sigma^2 \right) - \frac{1}{2}\, m\omega^2 \,. \end{aligned} \tag{19.38}$$

We also find:

$$f_{ij}(\Gamma, \Sigma) = \hbar \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \,, \qquad f^{ij}(\Gamma, \Sigma) = \frac{1}{\hbar} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \,, \tag{19.39}$$

which is of the required symplectic form. From the equations of motion, we find that the Hamiltonian is a constant of the motion. We put:

$$H(\Gamma, \Sigma) = \frac{\hbar^2}{2m}\, \frac{1}{4\Gamma} \Big\{ 1 + (\,4\,\Gamma\,\Sigma\,)^2 \Big\} + \frac{1}{2}\, m\omega^2\,\Gamma = E = \frac{1}{2}\, m\omega^2\, a^2 \,, \tag{19.40}$$

where a is the classical turning point. From the equations of motion, we have:

$$\begin{aligned} \hbar\,\ddot{\Gamma} &= \frac{4\hbar^2}{m} \left\{ \dot{\Gamma}\,\Sigma + \Gamma\,\dot{\Sigma} \,, \right\} \\ &= \frac{4\hbar}{m} \left\{ \left(\frac{4\hbar^2}{m} \right) \Gamma\,\Sigma^2 + \left(\frac{\hbar^2}{2m} \right) \left[\frac{1}{4\Gamma} - 4\,\Gamma\,\Sigma^2 \right] - \frac{1}{2}\, m\omega^2\,\Gamma \right\} \\ &= \frac{4\hbar}{m} \left\{ E - m\omega^2\,\Gamma \right\} \,, \end{aligned}$$

from which we find:

$$\ddot{\Gamma} + (2\,\omega)^2\,\Gamma = (2\,\omega)^2\,(a^2/2) \,, \tag{19.41}$$

the solution of which is:

$$\Gamma(t) = \left\{ a^2 + R^2\, \cos(2\omega t - \phi) \right\} / \, 2 \,, \tag{19.42}$$

where R and ϕ are to be fixed by the energy and initial conditions. Since

$$\dot{\Gamma}(t)/\omega = -R^2\, \sin(2\omega t - \phi) \,, \tag{19.43}$$

from Eq. (19.40), we find:

$$\begin{aligned} \frac{1}{2}\, m\omega^2\, a^2 &= \frac{\hbar^2}{2m}\, \frac{1}{4\Gamma} \Big\{ 1 + \big(\,4\,\Gamma\,\Sigma\,\big)^2 \Big\} + \frac{1}{2}\, m\omega^2\,\Gamma \\ &= \frac{\hbar^2}{2m}\, \frac{1}{4\Gamma} \Big\{ 1 + \big(m\dot{\Gamma}/\hbar\big)^2 \Big\} + \frac{1}{2}\, m\omega^2\,\Gamma \,, \end{aligned} \tag{19.44}$$

which can be written as:

$$a^2\,\Gamma^2 = \left(\dot{\Gamma}/\omega\right)^2 + \left(b^2/2\right)^2 + \Gamma^2\,,$$

where b is the oscillator parameter, $b = \sqrt{\hbar/(m\omega)}$. Substituting our solutions (19.42) and (19.43) into this expression gives:

$$R^4 = a^4 - b^4 > 0\,, \qquad \text{so that:} \qquad R^2 = \sqrt{a^4 - b^4}\,. \tag{19.45}$$

which fixes R^2 in terms of the energy and the oscillator parameter. It is useful to rewrite $\Gamma(t)$ and $\dot{\Gamma}(t)$ in the following way. We note from Eq. (19.42) that we can write:

$$\begin{aligned}
\Gamma(t) &= \left\{\, a^2 + R^2\,\cos(2\omega t - \phi)\right\}/\,2 \\
&= b^2\left\{\,\cosh(2r) + \sinh(2r)\,\cos(2\omega t - \phi)\right\}/\,2 \\
&= b^2\left\{\,\cosh^2(r) + \sinh^2(r) + 2\,\sinh(r)\cosh(r)\,\cos(2\omega t - \phi)\,\right\}/\,2 \\
&= b^2\left\{\,\cosh(r) + e^{+i(2\omega t-\phi)}\,\sinh(r)\,\right\}\left\{\,\cosh(r) + e^{-i(2\omega t-\phi)}\,\sinh(r)\,\right\}/\,2\,.
\end{aligned} \tag{19.46}$$

where

$$\cosh(2r) = (a/b)^2\,, \qquad \sinh(2r) = \sqrt{(a/b)^2 - 1}\,. \tag{19.47}$$

For $\Sigma(t)$, we find:

$$\Sigma(t) = \frac{m}{4\hbar}\,\frac{\dot{\Gamma}(t)}{\Gamma(t)} = \left(\frac{1}{2b^2}\right)\frac{-\sinh(2r)\,\sin(2\omega t - \phi)}{\cosh(2r) + \sinh(2r)\,\cos(2\omega t - \phi)}\,, \tag{19.48}$$

so that

$$\begin{aligned}
\frac{1}{4\Gamma(t)} - i\Sigma(t) &= \left(\frac{1}{2b^2}\right)\frac{1 + i\,\sinh(2r)\,\sin(2\omega t - \phi)}{\cosh(2r) + \sinh(2r)\,\cos(2\omega t - \phi)} \\
&= \left(\frac{1}{2b^2}\right)\frac{\cosh^2(r) - \sinh^2(r) + 2i\,\sinh(r)\cosh(r)\,\sin(2\omega t - \phi)}{\cosh^2(r) + \sinh^2(r) + 2\,\sinh(r)\cosh(r)\,\cos(2\omega t - \phi)} \\
&= \left(\frac{1}{2b^2}\right)\frac{\left\{\,\cosh(r) - e^{-i(2\omega t-\phi)}\,\sinh(r)\,\right\}\left\{\,\cosh(r) + e^{+i(2\omega t-\phi)}\,\sinh(r)\,\right\}}{\left\{\,\cosh(r) + e^{-i(2\omega t-\phi)}\,\sinh(r)\,\right\}\left\{\,\cosh(r) + e^{+i(2\omega t-\phi)}\,\sinh(r)\,\right\}} \\
&= \left(\frac{1}{2b^2}\right)\frac{\cosh(r) - e^{-i(2\omega t-\phi)}\,\sinh(r)}{\cosh(r) + e^{-i(2\omega t-\phi)}\,\sinh(r)}\,.
\end{aligned}$$

From (19.37) and (19.38), the Lagrangian is given by:

$$L(t) = -\frac{\hbar^2}{4m\,\Gamma(t)}\,, \tag{19.49}$$

so the phase angle is:

$$\begin{aligned}
\theta(t) &= \int_0^t L(t)\,\mathrm{d}t/\hbar = -\frac{\hbar}{4m}\int_0^t L(t)\,\mathrm{d}t\,\frac{\mathrm{d}t}{\Gamma(t)} = -\frac{1}{2}\int_{x_0}^{x}\frac{\mathrm{d}x}{\cosh(2r) + \sinh(2r)\,\cos(2x)} \\
&= -\frac{1}{2}\int_{x_0}^{x}\frac{\mathrm{d}x}{(\cosh(r) + \sinh(r))^2\,\cos^2(x) + (\cosh(r) - \sinh(r))^2\,\sin^2(x)} \\
&= -\frac{1}{2}\,\frac{1}{(\cosh(r) + \sinh(r))^2}\int_{x_0}^{x}\frac{\mathrm{d}x}{\cos^2(x)\,(\,1 + \beta^2\,\tan^2(x))}\,,
\end{aligned}$$

where we have put $x = \omega t - \phi/2$, so that $x_0 = -\phi/2$, and have defined β by:

$$\beta = \frac{\cosh(r) - \sinh(r)}{\cosh(r) + \sinh(r)}\,.$$

With the substitution $u = \beta\tan(x)$, we obtain:

$$-2\theta(t) = \int_{u_0}^{u} \frac{\mathrm{d}u}{1+u^2} = \tan^{-1}(u) - \tan^{-1}(u_0) = \tan^{-1}(K)\,, \tag{19.50}$$

where we have set K equal to:

$$K = \frac{u-u_0}{1+u\,u_0} = \frac{\beta\tan(x) - \beta\tan(x_0)}{1+\beta^2\,\tan(x)\tan(x_0)}\,. \tag{19.51}$$

Inverting expression (19.50), we find:

$$\tan(2\theta(t)) = -K$$
$$\frac{e^{2i\theta(t)} - e^{-2i\theta(t)}}{e^{2i\theta(t)} + e^{-2i\theta(t)}} = -iK$$

or, if we put $z = e^{2i\theta(t)}$, then

$$\frac{z - 1/z}{z + 1/z} = \frac{z^2-1}{z^2+1} = -iK$$

which has the solution:

$$\begin{aligned} z^2 &= \frac{1-iK}{1+iK} \\ &= \frac{1+\beta^2\,\tan(x)\tan(x_0) - i\,(\,\tan(x) - \tan(x_0)\,)}{1+\beta^2\,\tan(x)\tan(x_0) + i\,(\,\tan(x) - \tan(x_0)\,)} \\ &= \frac{a_+^2\,\cos(x)\cos(x_0) + a_-^2\,\sin(x)\sin(x_0) - i\,(\,\sin(x)\cos(x_0) - \cos(x)\sin(x_0)\,)}{a_+^2\,\cos(x)\cos(x_0) + a_-^2\,\sin(x)\sin(x_0) + i\,(\,\sin(x)\cos(x_0) - \cos(x)\sin(x_0)\,)}\,, \end{aligned}$$

where we have set:

$$a_+ = \cosh(r) + \sinh(r)\,, \qquad a_- = \cosh(r) - \sinh(r)\,,$$

so that $a_+a_- = 1$. So that z^2 can be written as:

$$\begin{aligned} e^{4i\theta(t)} = z^2 &= \frac{(\,a_+\cos(x) - ia_-\sin(x)\,)\,(\,a_+\cos(x_0) + ia_-\sin(x_0)\,)}{(\,a_+\cos(x) + ia_-\sin(x)\,)\,(\,a_+\cos(x_0) - ia_-\sin(x_0)\,)} \\ &= \frac{(\,\cosh(r)\,e^{-ix} + \sinh(r)\,e^{+ix}\,)\,(\,\cosh(r)\,e^{+ix_0} + \sinh(r)\,e^{-ix_0}\,)}{(\,\cosh(r)\,e^{+ix} + \sinh(r)\,e^{-ix}\,)\,(\,\cosh(r)\,e^{-ix_0} + \sinh(r)\,e^{+ix_0}\,)} \\ &= \frac{(\,\cosh(r) + \sinh(r)\,e^{+2ix}\,)\,(\,\cosh(r) + \sinh(r)\,e^{-2ix_0}\,)}{(\,\cosh(r) + \sinh(r)\,e^{-2ix}\,)\,(\,\cosh(r) + \sinh(r)\,e^{+2ix_0}\,)}\,e^{-2i(x-x_0)}\,. \end{aligned}$$

But since $2x = 2\omega t - \phi$, $2x_0 = -\phi$, and $2(x-x_0) = 2\omega t$, the normalization factor together with the time-dependent phase becomes:

$$\frac{e^{i\theta(t)}}{[\,2\pi\Gamma(t)\,]^{1/4}} = \frac{e^{i(\phi-\theta_0-2\omega t)/4}}{\pi^{1/4}\,[\,b\,(\cosh(r) + e^{-2i\omega t}\sinh(r))\,]^{1/2}} \tag{19.52}$$

where the phase θ_0 is given by:

$$\tan(\theta_0/2) = e^{-2r}\,\tan(\phi/2)\,. \tag{19.53}$$

Putting all this together, the variational wave function (19.35) for the harmonic oscillator is given by:

$$\psi(x,t) = \frac{\exp\Big\{-\Big(\frac{x^2}{2b^2}\Big)\,\frac{\cosh(r) - e^{i(2\omega t-\phi)}\,\sinh(r)}{\cosh(r) + e^{i(2\omega t-\phi)}\,\sinh(r)} + \frac{i}{4}(\phi - \theta_0 - 2\omega t)\Big\}}{\pi^{1/4}\,[\,b\,(\cosh(r) + e^{-i(2\omega t-\phi)}\sinh(r))\,]^{1/2}}\,, \tag{19.54}$$

in agreement with our previous result for a squeezed state, with a squeeze parameter r. Thus this solution is exact.

We note that our solutions for this parameterization yields periodic orbits, so that we can use them to find eigenvalues of the system. From our solutions, Eqs. (19.46) and (19.48), we see that all $(\Gamma(t), \Sigma(t))$ orbits in phase space are periodic, with period $T = \pi/\omega$, which is *half* the value of the classical orbit for a harmonic oscillator. We need to find the phase space integral $I(T)$ for one of these orbits. We find:

$$I(T) = \int_0^T \pi_i(t)\, \dot{y}^i(t)\, \frac{\mathrm{d}t}{2\pi\hbar} \tag{19.55}$$

But $\pi_i(t)\, \dot{y}^i(t) = E - L(t)$, so we get:

$$I(T) = \frac{ET}{2\pi\hbar} - \frac{\theta(T)}{2\pi} = \frac{E}{2\hbar\omega} - \frac{\theta(T)}{2\pi}\,. \tag{19.56}$$

In order to evaluate $\theta(T)$, we first note that $x(T) = \pi - \phi/2$, so that from (19.51)

$$K = \beta\, \frac{\tan(x(T)) - \tan(x_0)}{1 + \beta^2\, \tan(x(T))\tan(x_0)} = 0\,. \tag{19.57}$$

Then $\theta(T)$ is given by the solution to:

$$\tan(2\theta(T)) = 0\,. \tag{19.58}$$

After some consideration, the correct zero is given by $2\theta(T) = \pi$, so we find:

$$I(T) = \frac{E}{2\hbar\omega} - \frac{1}{4} \equiv n\,. \tag{19.59}$$

So the eigenvalues are given by:

$$E = \hbar\omega\left(2n + \frac{1}{2}\right), \tag{19.60}$$

with $n = 0, 1, 2, \ldots$. Eq. (19.60) is the exact result. We only get the *even* eigenvalues because we picked a trial wave function which was symmetric about the origin. Thus our results happen to be exact because we chose a Gaussian form for the wave function. Note however that the Gaussian form we selected *is not an eigen function*, nevertheless the Gaussian form gave the exact eigenvalues. In the next section, we look at an example which does not have a simple analytic solution.

19.5.2 The anharmonic oscillator

The Gaussian trial wave function for the harmonic oscillator in the last section happened to be an exact solution, and the variational method for bound states gave the exact answer. In this section, we study the anharmonic oscillator.

Let us scale the Lagrangian so that in appropriate units, it is given by:

$$L = \frac{1}{2}\dot{x}^2 - x^4\,, \tag{19.61}$$

Again, we take a simple Gaussian variational trial wave function ϕ of the form,

$$\phi(x; N, G, \Sigma) = \frac{1}{[\,2\pi\Gamma(t)\,]^{1/4}} \exp\Big\{ i\theta(t) - x^2\left[\,\frac{1}{4\,\Gamma(t)} - i\,\Sigma(t)\,\right]\Big\}\,, \tag{19.62}$$

where $\Gamma(t)$ and $\Sigma(t)$ are the time-dependent variational parameters. The energy is now given by:

$$E(\Gamma, \Sigma) = 2\,\Gamma\,\Sigma^2 + \frac{1}{8\,\Gamma} + 3\,\Gamma^2\,,$$

Table 19.1: The first five energies of the anharmonic oscillator computed using the time-dependent variational method compared to the exact results [?] and a SUSY-based variational method [?].

n	variational	exact	SUSY
0	0.6814	0.6680	0.6693
1	6.6980	4.6968	4.7133
2	14.7235	10.2443	9.3102
3	24.0625	16.7118	
4	34.4217	23.8900	

and the equations of motion are:

$$\begin{aligned} \dot{\Gamma} &= 4\,\Gamma\,\Sigma\,, \\ \dot{\Sigma} &= -2\Sigma^2 - 6\,\Gamma + \frac{1}{8\,\Gamma^2}\,. \end{aligned}$$

Changing variables to

$$\Gamma = \rho^2\,, \qquad \Sigma = \frac{p_\rho}{2\,\rho}\,,$$

the equations of motion become

$$\dot{\rho} = p_\rho\,, \qquad \dot{p}_\rho = \frac{1}{4\,\rho^3} - 12\,\rho^3\,.$$

We can further scale the variables so as to completely remove the constants from the equations of motion. If we let

$$\rho = \sqrt{2}\,x\,, \qquad p_\rho = \sqrt{2}\,y\,,$$

then the energy equation becomes:

$$E = y^2 + 12\,x^4 + \frac{1}{16\,x^2}\,,$$

The action integral I for this case is given by:

$$I = \frac{4}{2\pi}\int_{x_{\min}}^{x_{\max}} \sqrt{E - 12\,x^4 - \frac{1}{16\,x^2}}\,\mathrm{d}x\,. \tag{19.63}$$

The turning points now have to be found numerically. The results for the first five (even) energy levels are given in Table 19.1, where we have compared these results with the exact (numerical) results of Hioe and Montroll [?] and the results of a SUSY-based variational method [?]. Note that the results for the variational approximation are upper bounds on the energies. However these energies are not very accurate in this case, and indicate that our assumed Gaussian form of the wave function does not capture the dynamics of the anharmonic oscillator very well.

19.5.3 Time-dependent Hartree-Fock

References

[1] P. Kramer and M. Saraceno, *Geometry of the time-dependent variational principle in quantum mechanics*, number 140 in Lecture Notes in Physics (1981).

Chapter 20

Exactly solvable potential problems

In the past several years, there has been a much deeper understanding of why the one-dimensional Schrödinger is analytically solvable for certain potentials. The factorization method introduced by Schrödinger [1], and used in Section **??** for the coulomb potential, was known in 1940. Infeld a Hull [2] developed the factorization method more fully in 1951. It appears that Gendenshtein [3] was the first to discover the principle of shape invariance and the surprising relation of supersymmetry to the analytic solution of potential problems in quantum mechanics.

In this chapter, we follow the review work of F. Cooper, A. Khare, and U. Sukhatme [**?**, **?**].

20.1 Supersymmetric quantum mechanics

Here we formulate supersymmetry for a general potential in one-dimensional quantum mechanics. We apply the general method discussed in Section **??** for the harmonic oscillator. There we had found that ...

20.2 The hierarchy of Hamiltonians

Here we develop the method.

20.3 Shape invariance

And here we explain shape invariance, and give some examples.

References

[1] E. Schrödinger, "A method of determining quantum mechanical eigenvalues and eigenfunctions," Proc. Roy. Irish Acad. A **46**, 9 (1940).

[2] L. Infeld and T. E. Hull, "The factorization method," Rev. Mod. Phys. **23**, 21 (1951).

[3] L. E. Gendenshtein, "Derivation of the exact spectra of the Schrödinger equation by means of Supersymmetry," JETP Lett. **38**, 356 (1983).

Chapter 21

Angular momentum

In this chapter, we discuss the theory of angular momentum in quantum mechanics and applications of the theory to many practical problems. The relationship between group theory and the generators of the group are much simpler for the rotation group than the complete Galilean group we studied in Chapter 9 on symmetries. The use of angular momentum technology is particularly important in applications in atomic and nuclear physics. Unfortunately there is a lot of overhead to learn about before one can become reasonably knowledgeable in the field and a proficient calculator. But the effort is well worth it — with a little work, you too can become an "angular momentum technician!"

We start in this chapter with the eigenvalue problem for general angular momentum operators, followed by a discussion of spin one-half and spin one systems. We then derive the coordinate representation of orbital angular momentum wave functions. After defining parity and time-reversal operations on eigenvectors of angular momentum, we then discuss several classical descriptions of coordinate system rotations, followed by a discussion of how eigenvectors of angular momentum are related to each other in rotated systems. We then show how to couple two, three, and four angular momentum systems and introduce $3j$, $6j$, and $9j$ coupling and recoupling coefficients. We then define tensor operators and prove various theorems useful for calculations of angular momentum matrix elements, and end the chapter with several examples of interest from atomic and nuclear physics.

You will find in Appendix G, a presentation of Schwinger's harmonic oscillator theory of angular momentum. This method, which involves Boson algebra, is very useful for calculation of rotation matrices and Clebsch-Gordan coefficients, but is not necessary for a general understanding of how to use angular momentum technology. We include it as a special topic, and use it to derive some general formulas.

A delightful collection of early papers on the quantum theory of angular momentum, starting with original papers by Pauli and Wigner, can be found in Biedenharn and Van Dam [1]. We adopt here the notation and conventions of the *latest edition* of Edmonds[2], which has become one of the standard reference books in the field.

21.1 Eigenvectors of angular momentum

The Hermitian angular momentum operators J_i, $i = 1, 2, 3$, obey the algebra:

$$[\, J_i, J_j \,] = i\hbar\, \epsilon_{ijk} J_k \tag{21.1}$$

In this section, we prove the following theorem:

Theorem 40. *The eigenvalues and eigenvectors of the angular momentum operator obey the equations:*

$$\begin{aligned} J^2 |\, j, m \,\rangle &= \hbar^2\, j(j+1) |\, j, m \,\rangle\,, \\ J_z |\, j, m \,\rangle &= \hbar\, m |\, j, m \,\rangle\,, \\ J_\pm |\, j, m \,\rangle &= \hbar\, A(j, \mp m) |\, j, m \pm 1 \,\rangle\,, \end{aligned} \tag{21.2}$$

where $J_\pm = J_x \pm iJ_y$, *and*

$$A(j,m) = \sqrt{(j+m)(j-m+1)}\,, \qquad A(j,1\pm m) = A(j,\mp m)\,, \tag{21.3}$$

with

$$j = 0, 1/2, 1, 3/2, 2, \dots\,, \qquad -j \le m \le j\,.$$

Proof. It is easy to see that $J^2 = J_z^2 + J_y^2 + J_z^2$ commutes with J_z: $[J^2, J_z] = 0$. Of course, J^2 commutes with any other component of $\mathbf{J}$. Thus, we can simultaneously diagonalize J^2 and any component of $\mathbf{J}$, which we choose to be J_z. We write these eigenvectors as $|\,\lambda, m\,\rangle$. They satisfy:

$$J^2|\,\lambda, m\,\rangle = \hbar^2\,\lambda\,|\,\lambda, m\,\rangle\,,$$
$$J_z|\,\lambda, m\,\rangle = \hbar\, m\,|\,\lambda, m\,\rangle\,.$$

We now define operators, $J_\pm$ by linear combinations of J_x and J_y: $J_\pm = J_x \pm iJ_y$, with the properies:

$$J_\pm^\dagger = J_\mp\,, \qquad [J_z, J_\pm] = \pm\hbar\, J_\pm\,, \qquad [J_+, J_-] = 2\hbar\, J_z$$

The total angular momentum can be written in terms of $J_\pm$ and J_z in several ways. We have:

$$J^2 = \frac{1}{2}(J_-J_+ + J_+J_-) + J_z^2 = J_+J_- + J_z^2 - \hbar J_z = J_-J_+ + J_z^2 + \hbar J_z\,. \tag{21.4}$$

The ladder equations are found by considering,

$$J_z\,\{J_\pm|\,\lambda, m\,\rangle\} = (J_\pm J_z + [J_z, J_\pm])\,|\,\lambda, m\,\rangle = \hbar\,(m\pm 1)\,\{J_\pm|\,\lambda, m\,\rangle\}\,.$$

Therefore $J_\pm|\,\lambda, m\,\rangle$ is an eigenvector of J_z with eigenvalue $\hbar(m\pm 1)$. So we can write:

$$J_+|\,\lambda, m\,\rangle = \hbar\, B(\lambda, m)|\,\lambda, m+1\,\rangle\,, \tag{21.5}$$
$$J_-|\,\lambda, m\,\rangle = \hbar\, A(\lambda, m)|\,\lambda, m-1\,\rangle\,.$$

But since $J_- = J_+^\dagger$, it is easy to show that $B(\lambda, m) = A^*(\lambda, m+1)$.

Using (21.4), we find that m is bounded from above and below. We have:

$$\langle\,\lambda, m\,|\{J^2 - J_z^2\}|\,\lambda, m\,\rangle = \hbar^2\,(\lambda - m^2) = \frac{1}{2}\langle\,\lambda, m\,|(J_+^\dagger J_+ + J_-^\dagger J_-)|\,\lambda, m\,\rangle \ge 0\,.$$

So $0 \le m^2 \le \lambda$. Thus, for fixed $\lambda \ge 0$, m is bounded by: $-\sqrt{\lambda} \le m \le +\sqrt{\lambda}$. Thus there must be a maximum and a minimum m, which we call $m_{\max}$, and $m_{\min}$. This means that there must exist some ket, $|\,\lambda, m_{\max}\,\rangle$, such that:

$$J_+|\,\lambda, m_{\max}\,\rangle = 0\,,$$
$$\text{or,} \qquad J_-J_+|\,\lambda, m_{\max}\,\rangle = (J^2 - J_z^2 - \hbar J_z)|\,\lambda, m_{\max}\,\rangle$$
$$= \hbar^2(\lambda - m_{\max}^2 - m_{\max})|\,\lambda, m_{\max}\,\rangle = 0\,,$$

so $m_{\max}(m_{\max}+1) = \lambda$. Similarly, there must exist some other ket, $|\,\lambda, m_{\min}\,\rangle$ such that:

$$J_-|\,\lambda, m_{\min}\,\rangle = 0\,,$$
$$\text{or,} \qquad J_+J_-|\,\lambda, m_{\min}\,\rangle = (J^2 - J_z^2 + \hbar J_z)|\,\lambda, m_{\min}\,\rangle$$
$$= \hbar^2(\lambda - m_{\min}^2 + m_{\min})|\,\lambda, m_{\min}\,\rangle = 0\,,$$

so we find that $m_{\min}(m_{\min}-1) = \lambda$. Therefore we must have

$$m_{\max}(m_{\max}+1) = \lambda = m_{\min}(m_{\min}-1)\,,$$

Which means that either $m_{\min} = -m_{\max}$, which is possible, or $m_{\min} = m_{\max} + 1$, which is impossible! So we set $j = m_{\max} = -m_{\min}$, which defines j. Then $\lambda = m_{\max}(m_{\max}+1) = m_{\min}(m_{\min}-1) = j(j+1)$. Now we must be able to reach $|\,\lambda, m_{\max}\,\rangle$ from $|\,\lambda, m_{\min}\,\rangle$ by applying J_+ in unit steps. This means that $m_{\max} - m_{\min} = 2j = n$, where $n = 0, 1, 2, \ldots$ is an integer. So $j = n/2$ is *half-integral*.

We can find $A(j,m)$ and $B(j,m)$ by squaring the second of (21.5). We find:

$$\begin{aligned}\hbar^2|A(j,m)|^2\langle\, j,m-1\,|\,j,m-1\,\rangle &= \langle\, j,m\,|J_+J_-|\,j,m\,\rangle\,,\\ &= \langle\, j,m\,|(J^2 - J_z^2 + \hbar J_z)|\,j,m\,\rangle\,,\\ &= \hbar^2\{j(j+1) - m^2 + m\}\,,\\ &= \hbar^2(j+m)(j-m+1)\,.\end{aligned}$$

Taking $A(j,m)$ to be real (this is conventional), we find:

$$A(j,m) = \sqrt{(j+m)(j-m+1)}\,,$$

which also determines $B(j,m) = A(j,m+1)$. This completes the proof. □

Remark 29. Note that we used only the commutation properties of the components of angular momentum, and did not have to consider any representation of the angular momentum operators.

Remark 30. The appearance of half-integer quantum numbers for j is due to the fact that there exists a two-dimensional representation of the rotation group. We will discuss this connection in Section 21.2.4 below.

Remark 31. The eigenvectors of angular momentum $|\,j,m\,\rangle$ refer to a particular coordinate frame Σ, where we chose to find common eigenvectors of J^2 and J_z in that frame. We can also find common angular momentum eigenvectors of J^2 and $J_{z'}$, referred to some other frame Σ', which is rotated with respect to Σ. We write these eigenvectors as $|\,j,m\,\rangle'$. They have the same values for j and m, and are an equivalent description of the system, and so are related to the eigenvectors $|\,j,m\,\rangle$ by a unitary transformation. We find these unitary transformations in Section 21.3 below.

21.1.1 Spin

The spin operator $\mathbf{S}$ is a special case of the angular momentum operator. It may *not* have a coordinate representation. The possible eigenvalues for the magnitude of intrinsic spin are $s = 0, 1/2, 1, 3/2, \ldots$.

Spin one-half

The case when $s = 1/2$ is quite important in angular momentum theory, and we have discussed it in great detail in Chapter 15. We only point out here that the Pauli spin-1/2 matrices are a special case of the general angular momentum problem we discussed in the last section. Using the results of Theorem 40 for the case of $j = 1/2$, the matrix elements of the spin one-half angular momentum operator is given by:

$$\langle\, 1/2,m\,|\,(\,J_x + iJ_y\,)\,|\,1/2,m'\,\rangle = \hbar\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\,,\qquad \langle\, 1/2,m\,|\,(\,J_x - iJ_y\,)\,|\,1/2,m'\,\rangle = \hbar\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\,,$$

$$\langle\, 1/2,m\,|\,J_z\,|\,1/2,m'\,\rangle = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\,,$$

So the matrices for spin-1/2 can be written in terms of the Pauli matrices by writing: $\mathbf{S} = (\hbar/2)\,\boldsymbol{\sigma}$, where $\boldsymbol{\sigma} = \sigma_x\hat{\mathbf{x}} + \sigma_y\hat{\mathbf{y}} + \sigma_z\hat{\mathbf{z}}$ is a matrix of unit vectors, and where the Pauli matrices are given by:

$$\sigma_x = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\,,\qquad \sigma_y = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\,,\qquad \sigma_z = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\,. \tag{21.6}$$

The Pauli matrices are Hermitian, traceless matrices which obey the algebra:

$$\sigma_i\,\sigma_j+\sigma_j\,\sigma_i=2\,\delta_{ij}\,,\qquad \sigma_i\,\sigma_j-\sigma_j\,\sigma_i=2\,i\,\epsilon_{ijk}\,\sigma_k\,,$$
$$\text{or:}\qquad \sigma_i\,\sigma_j=\delta_{ij}+i\,\epsilon_{ijk}\,\sigma_k\,, \tag{21.7}$$

A spin one-half particle is fully described by a spinor $\chi(\theta,\phi)$ with two parameters of the form:

$$\chi(\theta,\phi)=\begin{pmatrix}e^{-i\phi/2}\cos(\theta/2)\\ e^{+i\phi/2}\sin(\theta/2)\end{pmatrix}\,, \tag{21.8}$$

where (θ,ϕ) is the direction of a unit vector $\hat{\mathbf{p}}$. $\chi(\theta,\phi)$ is an eigenvector of $\hat{\mathbf{p}}\cdot\boldsymbol{\sigma}$ with eigenvalue $+1$, i.e. spin-up in the $\hat{\mathbf{p}}$ direction. Here $\hat{\mathbf{p}}$ is called the **polarization** vector. The density matrix for spin one-half can be written in terms of just one unit vector $(\hat{\mathbf{p}})$ described by two polar angles (θ,ϕ):

$$\rho(\hat{\mathbf{p}})=\chi(\theta,\phi)\,\chi^\dagger(\theta,\phi)=\frac{1}{2}\,(\,1+\hat{\mathbf{p}}\cdot\boldsymbol{\sigma}\,)\,. \tag{21.9}$$

This result will be useful for describing a beam of spin one-half particles.

Spin one

The Deuteron has spin one. The spinor χ describing a spin one particle is a 3×1 matrix with three complex components. Since one of these is an overall phase, it takes eight real parameters to fully specify a spin-one spinor. In contrast, it takes only two real parameters to fully describe a spin one-half particle, as we found in the last section. The density matrix $\rho=\chi\,\chi^\dagger$ is a 3×3 Hermitian matrix and so requires nine basis matrices to describe it, one of which can be the unit matrix. That leaves eight more independent matrices which are needed. It is traditional to choose these to be combinations of the spin-one angular momentum matrices. From the results of Theorem 40, the matrix elements for the $j=1$ angular momentum operator is given by:

$$\langle\,1,m\,|\,(\,J_x+iJ_y\,)\,|\,1,m'\,\rangle=\hbar\begin{pmatrix}0&\sqrt{2}&0\\0&0&\sqrt{2}\\0&0&0\end{pmatrix}\,,\qquad \langle\,1,m\,|\,(\,J_x-iJ_y\,)\,|\,1,m'\,\rangle=\hbar\begin{pmatrix}0&0&0\\\sqrt{2}&0&0\\0&\sqrt{2}&0\end{pmatrix}\,,$$
$$\langle\,1,m\,|\,J_z\,|\,1,m'\,\rangle=\hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}\,,$$

So let us put $\mathbf{J}=\hbar\,\mathbf{S}$, where

$$S_x=\frac{1}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix}\,,\qquad S_y=\frac{1}{\sqrt{2}}\begin{pmatrix}0&-i&0\\i&0&-i\\0&i&0\end{pmatrix}\,,\qquad S_z=\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}\,. \tag{21.10}$$

The spin one angular momentum matrices obey the commutation relations: $[\,S_i,S_j\,]=i\,\epsilon_{ijk}S_k$. Also they are Hermitian, $S_i^\dagger=S_i$, and traceless: $\mathrm{Tr}[\,S_i\,]=0$. They also obey $\mathrm{Tr}[\,S_i^2\,]=2$ and $\mathrm{Tr}[\,S_iS_j\,]=0$. An additional five independent matrices can be constructed by the traceless symmetric matrix of Hermitian matrices S_{ij}, defined by:

$$S_{ij}=\frac{1}{2}\big(\,S_iS_j+S_jS_i\,\big)-\frac{1}{3}\,\mathbf{S}\cdot\mathbf{S}\,,\qquad S_{ij}^\dagger=S_{ij}\,. \tag{21.11}$$

We also note here that $\mathrm{Tr}[\,S_{ij}\,]=0$ for all values of i and j. So then the density matrix for spin one particles can be written as:

$$\rho=\frac{1}{3}\,\big(\,1+\mathbf{P}\cdot\mathbf{S}+\sum_{ij}T_{ij}\,S_{ij}\,\big)\,, \tag{21.12}$$

and where $\mathbf{P}$ is a real vector with three components and T_{ij} a real symmetric traceless 3×3 matrix with five components. So P_i and T_{ij} provide *eight* independent quantities that are needed to fully describe a beam of spin one particles.

Exercise 47. Find all independent matrix components of S_{ij}. Find all values of $\mathrm{Tr}[\,S_i\,S_{jk}\,]$ and $\mathrm{Tr}[\,S_{ij}\,S_{kl}\,]$. Use these results to find $\mathrm{Tr}[\,\rho\,S_i\,]$ and $\mathrm{Tr}[\,\rho\,S_{ij}\,]$ in terms of P_i and T_{ij}.

Exercise 48. Show that for spin one, the density matrix is idempotent: $\rho^2 = \rho$. Find any restrictions this places on the values of P_i and T_{ij}.

21.1.2 Orbital angular momentum

The orbital angular momentum for a single particle is defined as:

$$\mathbf{L} = \mathbf{R} \times \mathbf{P}\,, \tag{21.13}$$

where $\mathbf{R}$ and $\mathbf{P}$ are operators for the position and momentum of the particle, and obey the commutation rules: $[\,X_i, P_i\,] = i\hbar\,\delta_{ij}$. Then it is easy to show that:

$$[\,L_i, L_j\,] = i\hbar\,\epsilon_{ijk} L_k\,, \tag{21.14}$$

as required for an angular momentum operator. Defining as before $L_\pm = L_x \pm i\,L_y$, we write the eigenvalues and eigenvectors for orbital angular momentum as:

$$\begin{aligned} L^2\,|\,\ell, m\,\rangle &= \hbar^2\,\ell(\ell+1)\,|\,\ell, m\,\rangle\,, \\ L_z\,|\,\ell, m\,\rangle &= \hbar\,m\,|\,\ell, m\,\rangle\,, \\ L_\pm\,|\,\ell, m\,\rangle &= \hbar\,A(\ell, \mp m)\,|\,\ell, m \pm 1\,\rangle\,, \end{aligned} \tag{21.15}$$

for $-\ell \le m \le +\ell$, and $\ell = 0, 1, 2, \ldots$. We will show below that ℓ has only integer values. We label eigenvectors of spherical coordinates by $|\,\hat{\mathbf{r}}\,\rangle \mapsto |\,\theta, \phi\,\rangle$, and define:

$$Y_{\ell,m}(\hat{\mathbf{r}}) = \langle\,\hat{\mathbf{r}}\,|\,\ell, m\,\rangle = \langle\,\theta, \phi\,|\,\ell, m\,\rangle = Y_{\ell,m}(\theta, \phi)\,. \tag{21.16}$$

In the coordinate representation, $\tilde{\mathbf{L}}$ is a differential operator acting on functions:

$$\tilde{\mathbf{L}}\,Y_{\ell,m}(\theta, \phi) = \langle\,\hat{\mathbf{r}}\,|\,\mathbf{L}\,|\,\ell, m\,\rangle = \frac{\hbar}{i}\,\hat{\mathbf{r}} \times \boldsymbol{\nabla}\,Y_{\ell,m}(\theta, \phi)\,, \tag{21.17}$$

We can easily work out the orbital angular momentum in spherical coordinates. Using

$$x = r\sin\theta\cos\phi\,, \qquad y = r\sin\theta\sin\phi\,, \qquad z = r\cos\theta\,, \tag{21.18}$$

with spherical unit vectors defined by:

$$\begin{aligned} \hat{\mathbf{r}} &= \sin\theta\cos\phi\,\hat{\mathbf{x}} + \sin\theta\sin\phi\,\hat{\mathbf{y}} + \cos\theta\,\hat{\mathbf{z}} \\ \hat{\boldsymbol{\phi}} &= -\sin\phi\,\hat{\mathbf{x}} + \cos\phi\,\hat{\mathbf{y}} \\ \hat{\boldsymbol{\theta}} &= \cos\theta\cos\phi\,\hat{\mathbf{x}} + \cos\theta\sin\phi\,\hat{\mathbf{y}} - \sin\theta\,\hat{\mathbf{z}}\,, \end{aligned} \tag{21.19}$$

we find that the gradient operator is given by:

$$\boldsymbol{\nabla} = \hat{\mathbf{r}}\,\frac{\partial}{\partial r} + \hat{\boldsymbol{\phi}}\,\frac{1}{r\sin\theta}\,\frac{\partial}{\partial\phi} + \hat{\boldsymbol{\theta}}\,\frac{1}{r}\,\frac{\partial}{\partial\theta}\,. \tag{21.20}$$

So in the coordinate representation, the vector angular momentum operator is given by:

$$\tilde{\mathbf{L}} = \frac{\hbar}{i}\,\mathbf{r} \times \boldsymbol{\nabla} = \frac{\hbar}{i}\left\{\hat{\mathbf{r}} \times \hat{\boldsymbol{\phi}}\,\frac{1}{\sin\theta}\,\frac{\partial}{\partial\phi} + \hat{\mathbf{r}} \times \hat{\boldsymbol{\theta}}\,\frac{\partial}{\partial\theta}\right\} = \frac{\hbar}{i}\left\{-\hat{\boldsymbol{\theta}}\,\frac{1}{\sin\theta}\,\frac{\partial}{\partial\phi} + \hat{\boldsymbol{\phi}}\,\frac{\partial}{\partial\theta}\right\}, \tag{21.21}$$

which is independent of the radial coordinate r. Components are given by:

$$\begin{aligned}\tilde{L}_x &= \frac{\hbar}{i}\left\{-\sin\phi\,\frac{\partial}{\partial\theta} - \frac{\cos\phi}{\tan\theta}\,\frac{\partial}{\partial\phi}\right\},\\ \tilde{L}_y &= \frac{\hbar}{i}\left\{+\cos\phi\,\frac{\partial}{\partial\theta} - \frac{\sin\phi}{\tan\theta}\,\frac{\partial}{\partial\phi}\right\},\\ \tilde{L}_z &= \frac{\hbar}{i}\left\{\frac{\partial}{\partial\phi}\right\},\end{aligned} \tag{21.22}$$

from which we get:

$$\tilde{L}_\pm = L_x \pm i\,L_y = \frac{\hbar}{i}\,e^{\pm i\phi}\left\{\pm i\,\frac{\partial}{\partial\theta} - \frac{1}{\tan\theta}\,\frac{\partial}{\partial\phi}\right\}, \tag{21.23}$$

and so

$$\tilde{L}^2 = \frac{1}{2}(L_+L_- + L_-L_+) + L_z^2 = -\hbar^2\Big\{\frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\Big(\sin\theta\,\frac{\partial}{\partial\theta}\Big) + \frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\phi^2}\Big\}, \tag{21.24}$$

Single valued eigenfunctions of L^2 and L_z are the spherical harmonics, $Y_{\ell m}(\theta,\phi)$, given by the solution of the equations,

$$\begin{aligned}-\hbar^2\left\{\frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\phi^2}\right\}Y_{\ell,m}(\theta,\phi) &= \hbar^2\,\ell(\ell+1)\,Y_{\ell,m}(\theta,\phi)\,,\\ \frac{\hbar}{i}\left\{\frac{\partial}{\partial\phi}\right\}Y_{\ell,m}(\theta,\phi) &= \hbar\,m\,Y_{\ell m}(\theta,\phi)\,,\\ \frac{\hbar}{i}\,e^{\pm i\phi}\left\{\pm i\,\frac{\partial}{\partial\theta} - \frac{1}{\tan\theta}\,\frac{\partial}{\partial\phi}\right\}Y_{\ell,m}(\theta,\phi) &= \hbar\,A(\ell,\mp m)\,Y_{\ell,m\pm1}(\theta,\phi)\,,\end{aligned} \tag{21.25}$$

where $\ell = 0, 1, 2, \ldots$, with $-\ell \le m \le \ell$, and $A(\ell,m) = \sqrt{(\ell+m)(\ell-m+1)}$. Note that the eigenvalues of the orbital angular momentum operator are integers. The half-integers eigenvalues of general angular momentum operators are missing from the eigenvalue spectra. This is because wave functions in coordinate space must be single valued.

Definition 31 (spherical harmonics)**.** We define spherical harmonics by:

$$Y_{\ell,m}(\theta,\phi) = \begin{cases}\sqrt{\dfrac{2\ell+1}{4\pi}\,\dfrac{(\ell-m)!}{(\ell+m)!}}\,(-)^m\,e^{im\phi}\,P_\ell^m(\cos\theta)\,, & \text{for } m \ge 0,\\ (-)^m\,Y^*_{\ell,-m}(\theta,\phi)\,, & \text{for } m < 0.\end{cases} \tag{21.26}$$

where $P_\ell^m(\cos\theta)$ are the associated Legendre polynomials which are *real* and depend only on $|m|$. This is Condon and Shortly's definition [3], which is the same as Edmonds [2][pages 19–25] and is now standard.

The spherical harmonics defined here have the properites:

- The spherical harmonics are orthonormal and complete:

$$\int Y^*_{\ell m}(\Omega)\,Y_{\ell' m'}(\Omega)\,\mathrm{d}\Omega = \delta_{\ell,\ell'}\delta_{m,m'}\,, \qquad \sum_{\ell m} Y^*_{\ell m}(\Omega)\,Y_{\ell m}(\Omega') = \delta(\Omega-\Omega')\,,$$

  where $\mathrm{d}\Omega = \mathrm{d}(\cos\theta)\,\mathrm{d}\phi$.

- Under complex conjugation,

$$Y^*_{\ell,m}(\theta,\phi) = (-)^m\,Y_{\ell,-m}(\theta,\phi)\,. \tag{21.27}$$

- Under space inversion:

$$Y_{\ell,m}(\pi-\theta,\phi+\pi) = (-)^\ell\,Y_{\ell,m}(\theta,\phi)\,. \tag{21.28}$$

- We also note that since $P_\ell^m(\cos\theta)$ is real,

$$Y_{\ell,m}(\theta,-\phi) = Y_{\ell,m}(\theta,2\pi-\phi) = Y^*_{\ell,m}(\theta,\phi)\,. \tag{21.29}$$

- At $\theta = 0$, $\cos\theta = 1$, $P_\ell^m(1) = \delta_{m,0}$ so that:

$$Y_{\ell,m}(0,\phi) = \sqrt{\frac{2\ell+1}{4\pi}}\,\delta_{m,0}\,, \tag{21.30}$$

  independent of ϕ.

Other properties of the spherical harmonics can be found in Edmonds [2] and other reference books. It us useful to know the first few spherical harmonics. These are:

$$\begin{aligned} Y_{0,0}(\theta,\phi) &= \sqrt{\frac{1}{4\pi}}\,, \qquad Y_{1,0}(\theta,\phi) = \sqrt{\frac{3}{4\pi}}\,\cos\theta\,, \qquad Y_{1,\pm1}(\theta,\phi) = \mp\sqrt{\frac{3}{8\pi}}\,\sin\theta\,e^{\pm i\phi}\,, \\ Y_{2,0}(\theta,\phi) &= \sqrt{\frac{5}{16\pi}}\,(\,2\cos^2\theta - \sin^2\theta\,)\,, \qquad Y_{2,\pm1}(\theta,\phi) = \mp\sqrt{\frac{15}{8\pi}}\,\cos\theta\,\sin\theta\,e^{\pm i\phi}\,, \\ Y_{2,\pm2}(\theta,\phi) &= \sqrt{\frac{15}{32\pi}}\,\sin^2\theta\,e^{\pm 2i\phi}\,. \end{aligned} \tag{21.31}$$

Definition 32 (Reduced spherical harmonics)**.** Sometimes it is useful to get rid of factors and define reduced spherical harmonics (Racah [4]) $C_{\ell,m}(\theta,\phi)$ by:

$$C_{\ell,m}(\theta,\phi) = \sqrt{\frac{4\pi}{2\ell+1}}\,Y_{\ell,m}(\theta,\phi)\,. \tag{21.32}$$

Remark 32. The orbital angular momentum states for $\ell = 0,1,2,3,4,\ldots$ are often referred to as $s,p,d,f,g,\ldots$ states.

21.1.3 Kinetic energy operator

In this section, we relate the kinetic energy operator for a single particle to orbital angular momentum. We first note that in spherical coordinates, coordinate representations of operators should be defined to correspond to the usual coordinate transformation from Cartesian to spherical coordinates. That is:

$$\hat{\mathbf{R}}\cdot\mathbf{P} = \frac{\mathbf{R}}{R}\cdot\mathbf{P} \qquad \text{so that} \qquad \langle\,\mathbf{r}\,|\,\hat{\mathbf{R}}\cdot\mathbf{P}\,|\,\psi\,\rangle = \frac{\hbar}{i}\,\frac{\mathbf{r}}{r}\cdot\boldsymbol{\nabla}\,\psi(\mathbf{r}) = \frac{\hbar}{i}\,\frac{\partial\psi(\mathbf{r})}{\partial r}\,. \tag{21.33}$$

This means that we should require the operator relation:

$$[\,R,\hat{\mathbf{R}}\cdot\mathbf{P}\,] = i\hbar\,. \tag{21.34}$$

However in spherical coordinates, $\hat{\mathbf{R}}\cdot\mathbf{P}$, is *not* Hermitian. However, we see that we can fix this by defining an operator P_r by:

$$P_r = \frac{1}{R}\,[\,\mathbf{R}\cdot\mathbf{P} - i\hbar\,]\,, \qquad \text{so that} \qquad \langle\,\mathbf{r}\,|\,P_r\,|\,\psi\,\rangle = \frac{\hbar}{i}\Big[\,\frac{\partial}{\partial r} + \frac{1}{r}\,\Big]\,\psi(\mathbf{r})\,. \tag{21.35}$$

Let us now show that P_r *is* Hermitian. We first note that $[\,\mathbf{R}\cdot\mathbf{P},\mathbf{P}\cdot\mathbf{R}\,] = 3i\hbar$ and that

$$[\,\mathbf{R}\cdot\mathbf{P},\frac{1}{R}\,] = \frac{1}{R}\,[\,R,\mathbf{R}\cdot\mathbf{P}\,]\,\frac{1}{R} = \frac{X_j}{R}\,[\,R,P_j\,]\,\frac{1}{R} = i\hbar\,\frac{X_j}{R}\,\frac{X_j}{R}\,\frac{1}{R} = \frac{i\hbar}{R}\,, \tag{21.36}$$

so that:

$$P_r^\dagger = \left[\, \mathbf{P}\cdot\mathbf{R} + i\hbar \,\right] \frac{1}{R} = \left[\, \mathbf{R}\cdot\mathbf{P} - 2i\hbar \,\right] \frac{1}{R} = \frac{1}{R}\,\mathbf{R}\cdot\mathbf{P} + \left[\, \mathbf{R}\cdot\mathbf{P}, \frac{1}{R} \,\right] - \frac{2i\hbar}{R} = \frac{1}{R}\left[\, \mathbf{R}\cdot\mathbf{P} - i\hbar \,\right] = P_r \,. \quad (21.37)$$

This miracle happens only in spherical coordinates, and is due to the factor of r^2 in the radial measure. We also have the commutation relation:

$$\left[\, R, P_r \,\right] = \left[\, R, \frac{1}{R}\mathbf{R}\cdot\mathbf{P} \,\right] = \frac{X_i}{R}\left[\, R, P_i \,\right] = i\hbar \,. \quad (21.38)$$

The square of the radial momentum operator is given by:

$$\begin{aligned} P_r^2 &= \left[\, \hat{\mathbf{R}}\cdot\mathbf{P} - \frac{i\hbar}{R} \,\right]^2 = (\hat{\mathbf{R}}\cdot\mathbf{P})\cdot(\hat{\mathbf{R}}\cdot\mathbf{P}) - \frac{i\hbar}{R}\,(\hat{\mathbf{R}}\cdot\mathbf{P}) - (\hat{\mathbf{R}}\cdot\mathbf{P})\,\frac{i\hbar}{R} - \frac{1}{R^2} \\ &= (\hat{\mathbf{R}}\cdot\mathbf{P})\,(\hat{\mathbf{R}}\cdot\mathbf{P}) - \frac{2i\hbar}{R}\,(\hat{\mathbf{R}}\cdot\mathbf{P}) = \frac{1}{R^2}\left[\, (\mathbf{R}\cdot\mathbf{P})\,(\mathbf{R}\cdot\mathbf{P}) - i\hbar\,(\mathbf{R}\cdot\mathbf{P}) \,\right] \end{aligned} \quad (21.39)$$

$$\text{so} \qquad \langle\, \mathbf{r} \,|\, P_r^2 \,|\, \psi \,\rangle = -\frac{\hbar}{i}\left[\, \frac{\partial^2}{\partial r^2} + \frac{2}{r}\,\frac{\partial}{\partial r} \,\right] \psi(\mathbf{r}) \,,$$

which we recognize as the radial part of the Laplacian operator. The kinetic energy operator can now be written in terms of P_r and the square of the angular momentum operator L^2. We notice that:

$$\begin{aligned} L^2 &= (\mathbf{R}\times\mathbf{P})\cdot(\mathbf{R}\times\mathbf{P}) = \mathbf{R}\cdot(\mathbf{P}\times(\mathbf{R}\times\mathbf{P})) \\ &= X_i P_j X_i P_j - X_i P_j X_j P_i = R^2\,P^2 - i\hbar\,(\mathbf{R}\cdot\mathbf{P}) - (\mathbf{R}\cdot\mathbf{P})\,(\mathbf{R}\cdot\mathbf{P}) + 2i\hbar\,(\mathbf{R}\cdot\mathbf{P}) \\ &= R^2\,P^2 - \left[\, (\mathbf{R}\cdot\mathbf{P})\,(\mathbf{R}\cdot\mathbf{P}) - i\hbar\,(\mathbf{R}\cdot\mathbf{P}) \,\right] = R^2\left\{\, P^2 - P_r^2 \,\right\} . \end{aligned} \quad (21.40)$$

So $P^2 = P_r^2 + L^2/R^2$, which is just a statement of what the Laplacian looks like in spherical coordinates. So the kinetic energy operator becomes:

$$T = \frac{P^2}{2m} = \frac{P_r^2}{2m} + \frac{L^2}{2m\,R^2} \,. \quad (21.41)$$

We will have occasion to use this definition of a radial momentum operator P_r when we discuss reduced matrix elements of the linear momentum tensor operator in Section 21.5.2, and in the operator factorization methods of Section 22.3.4.

21.1.4 Parity and Time reversal

We discussed the effects of parity and time reversal transformations on the generators of Galilean transformations, including the angular momentum generator, in Chapter 9. We study the effect of these transformations on angular momentum states in this section.

Parity

For parity, we found in Section 9.7.1 that $\mathcal{P}$ is linear and unitary, with eigenvalues of unit magnitude, and has the following effects on the angular momentum, position, and linear momentum operators:

$$\begin{aligned} \mathcal{P}^{-1}\,\mathbf{X}\,\mathcal{P} &= -\mathbf{X} \,, \\ \mathcal{P}^{-1}\,\mathbf{P}\,\mathcal{P} &= -\mathbf{P} \,, \\ \mathcal{P}^{-1}\,\mathbf{J}\,\mathcal{P} &= \mathbf{J} \,. \end{aligned} \quad (21.42)$$

We also found that $\mathcal{P}^{-1} = \mathcal{P}^\dagger = \mathcal{P}$. So under parity, we can take:

$$\mathcal{P}\,|\,\mathbf{x}\,\rangle = |\,-\mathbf{x}\,\rangle \,, \qquad \mathcal{P}\,|\,\mathbf{p}\,\rangle = |\,-\mathbf{p}\,\rangle \,. \quad (21.43)$$

The angular momentum operator does not change under parity, so $\mathcal{P}$ operating on a state of angular momentum $|\,jm\,\rangle$ can only result in a phase. If there is a coordinate representation of the angular momentum eigenstate, we can write:

$$\begin{aligned}\langle\,\hat{\mathbf{r}}\,|\,\mathcal{P}\,|\,\ell,m\,\rangle &= \langle\,\mathcal{P}^\dagger\,\hat{\mathbf{r}}\,|\,\ell,m\,\rangle = \langle\,\mathcal{P}\,\hat{\mathbf{r}}\,|\,\ell,m\,\rangle = \langle\,-\hat{\mathbf{r}}\,|\,\ell,m\,\rangle \\ &= Y_{\ell,m}(\pi-\theta,\phi+\pi) = (-)^\ell\,Y_{\ell,m}(\theta,\phi) = (-)^\ell\,\langle\,\mathbf{x}\,|\,\ell,m\,\rangle\,,\end{aligned}$$

where we have used (21.28). Therefore:

$$\mathcal{P}\,|\,\ell,m\,\rangle = (-)^\ell\,|\,\ell,m\,\rangle\,. \tag{21.44}$$

For spin 1/2 states, the parity operator must be the unit matrix. The phase is generally taken to be unity, so that:

$$\mathcal{P}\,|\,1/2,m\,\rangle = |\,1/2,m\,\rangle\,. \tag{21.45}$$

So parity has different results on orbital and spin eigenvectors.

Time reversal

For time reversal, we found in Section 9.7.2 that $\mathcal{T}$ is anti-linear and anti-unitary, $\mathcal{T}^{-1}i\mathcal{T} = -i$ with eigenvalues of unit magnitude, and has the following effects on the angular momentum, position, and linear momentum operators:

$$\begin{aligned}\mathcal{T}^{-1}\,\mathbf{X}\,\mathcal{T} &= \mathbf{X}\,,\\ \mathcal{T}^{-1}\,\mathbf{P}\,\mathcal{T} &= -\mathbf{P}\,,\\ \mathcal{T}^{-1}\,\mathbf{J}\,\mathcal{T} &= -\mathbf{J}\,.\end{aligned} \tag{21.46}$$

Under time-reversal,

$$\mathcal{T}\,|\,\mathbf{x}\,\rangle = |\,\mathbf{x}\,\rangle\,,\qquad \mathcal{T}\,|\,\mathbf{p}\,\rangle = |\,-\mathbf{p}\,\rangle\,. \tag{21.47}$$

The angular momentum operator reverses sign under time reversal, so $\mathcal{T}$ operating on a state of angular momentum can only result in a phase. Because of the anti-unitary property, the commutation relations for angular momentum are invariant under time reversal. However since $\mathcal{T}\,J^2\,\mathcal{T}^{-1} = J^2$, $\mathcal{T}\,J_z\,\mathcal{T}^{-1} = -J_z$, and $\mathcal{T}\,J_\pm\,\mathcal{T}^{-1} = -J_\mp$, operating on the eigenvalue equations (21.2) by $\mathcal{T}$ gives:

$$\begin{aligned}J^2\left\{\,\mathcal{T}\,|\,j,m\,\rangle\,\right\} &= \hbar^2\,j(j+1)\left\{\,\mathcal{T}\,|\,j,m\,\rangle\,\right\},\\ J_z\left\{\,\mathcal{T}\,|\,j,m\,\rangle\,\right\} &= -\hbar\,m\left\{\,\mathcal{T}\,|\,j,m\,\rangle\,\right\},\\ J_\mp\left\{\,\mathcal{T}\,|\,j,m\,\rangle\,\right\} &= -\,A(j,\mp m)\left\{\,\mathcal{T}\,|\,j,m\,\rangle\,\right\}.\end{aligned} \tag{21.48}$$

These equations have the solution:

$$\mathcal{T}\,|\,j,m\,\rangle = (-)^{j+m}\,|\,j,-m\,\rangle\,. \tag{21.49}$$

Here we have introduced an arbitrary phase $(-)^j$ so that for half-integer values of j, the operation of parity will produce a sign, not a complex number. Let us investigate time reversal on both spin-1/2 and integer values of j.

For spin-1/2 states, in a 2×2 matrix representation, we require:

$$\mathcal{T}^{-1}\,\sigma_i\,\mathcal{T} = -\sigma_i\,, \tag{21.50}$$

for $i = 1,2,3$. Now we know that σ_2 changes the sign of any σ_i, but it also takes the complex conjugate, which we do not want in this case. So for spin 1/2, we take the following matrix representation of the time reversal operator:

$$\mathcal{T} = i\,\sigma_2\,\mathcal{K} = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\mathcal{K}\,, \tag{21.51}$$

where $\mathcal{K}$ is a complex conjugate operator acting on functions. This makes $\mathcal{T}$ anti-linear and anti-unitary. Now since $(i\sigma_y)\sigma_x(i\sigma_y) = \sigma_x$, $(i\sigma_y)\sigma_y(i\sigma_y) = -\sigma_y$, and $(i\sigma_y)\sigma_z(i\sigma_y) = \sigma_z$, and recalling that σ_x and σ_z are real, whereas $\sigma_y^* = -\sigma_y$, so that:

$$\mathcal{T}^{-1}\,\sigma_i\,\mathcal{T} = -\sigma_i\,, \tag{21.52}$$

as required. Now the matrix representation of $\mathcal{T}$ on spinor states have the effect:

$$\mathcal{T}\,|\,1/2,1/2\,\rangle = i\,\sigma_2\,\mathcal{K}\,|\,1/2,1/2\,\rangle = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\mathcal{K}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -|\,1/2,-1/2\,\rangle\,.$$
$$\mathcal{T}\,|\,1/2,-1/2\,\rangle = i\,\sigma_2\,\mathcal{K}\,|\,1/2,-1/2\,\rangle = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\mathcal{K}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = +\begin{pmatrix} 1 \\ 0 \end{pmatrix} = +|\,1/2,+1/2\,\rangle\,,$$

so that

$$\mathcal{T}\,|\,1/2,m\,\rangle = (-)^{1/2+m}\,|\,1/2,-m\,\rangle\,, \tag{21.53}$$

in agreement with (21.49).

Exercise 49. For the spin $\mathcal{T}$ operator defined in Eq. (21.51), show that:

$$\mathcal{T}^{-1} = \mathcal{T}^\dagger = \mathcal{T}\,. \tag{21.54}$$

For integer values of the angular momentum, there is a coordinate representation of the angular momentum vector. If we choose

$$\langle\,\hat{\mathbf{r}}\,|\,\ell,m\,\rangle = Y_{\ell,m}(\theta,\phi)\,, \tag{21.55}$$

then we can write:

$$\begin{aligned}\langle\,\hat{\mathbf{r}}\,|\,\mathcal{T}\,|\,\ell,m\,\rangle &= \langle\,\mathcal{T}^\dagger\,\hat{\mathbf{r}}\,|\,\ell,m\,\rangle^* = \langle\,\mathcal{T}\,\hat{\mathbf{r}}\,|\,\ell,m\,\rangle^* = \langle\,\hat{\mathbf{r}}\,|\,\ell,m\,\rangle^* \\ &= Y^*_{\ell,m}(\theta,\phi) = (-)^m\,Y_{\ell,-m}(\theta,\phi) = (-)^m\,\langle\,\hat{\mathbf{r}}\,|\,\ell,-m\,\rangle\,.\end{aligned}$$

So we conclude that:

$$\mathcal{T}\,|\,\ell,m\,\rangle = (-)^m\,|\,\ell,-m\,\rangle\,, \tag{21.56}$$

which does *not* agree with (21.49). However if we choose:

$$\langle\,\hat{\mathbf{r}}\,|\,\ell,m\,\rangle = i^\ell\,Y_{\ell,m}(\theta,\phi)\,, \tag{21.57}$$

then

$$\begin{aligned}\langle\,\hat{\mathbf{r}}\,|\,\mathcal{T}\,|\,\ell,m\,\rangle &= \langle\,\mathcal{T}^\dagger\,\hat{\mathbf{r}}\,|\,\ell,m\,\rangle^* = \langle\,\mathcal{T}\,\hat{\mathbf{r}}\,|\,\ell,m\,\rangle^* = \langle\,\hat{\mathbf{r}}\,|\,\ell,m\,\rangle^* \\ &= \left[\,i^\ell\,Y_{\ell,m}(\theta,\phi)\,\right]^* = (-)^{\ell+m}\,Y_{\ell,-m}(\theta,\phi) = (-)^{\ell+m}\,\langle\,\hat{\mathbf{r}}\,|\,\ell,-m\,\rangle\,.\end{aligned}$$

which gives:

$$\mathcal{T}\,|\,\ell,m\,\rangle = (-)^{\ell+m}\,|\,\ell,-m\,\rangle\,, \tag{21.58}$$

which *does* agree with (21.49). We will see in Section 21.4 that when orbital and spin eigenvectors are coupled together by a Clebsch-Gordan coefficient, the operation of time reversal on the coupled state is preserved if we choose the spherical functions defined in Eq. (21.57). However, Eq. (21.55) is generally used in the literature.

21.2 Rotation of coordinate frames

A fixed point P in space, described by Euclidean coordinates (x,y,z) and (x',y',z') in two frames Σ and Σ', are related to each other by a rotation if lengths and angles are preserved. The same point in space is related to coordinates in these two systems by a linear orthogonal transformation of the form: $x'_i = R_{ij}\,x_j$, with $R_{ij}R_{ik} = \delta_{jk}$. Proper transformations which preserve orientation of the coordinate system are those with $\det[\,R\,] = +1$. The set of all orthogonal rotation matrices R form a group, called $SO(3)$, since:

1. The product RR' of any two group elements is another group element R''.
2. Matrix multiplication is associative: $(RR')R'' = R(R'R'')$.
3. There is a unique identity element $I = \delta_{ij}$, such that $I\,R = R$ for all R in the group, and
4. For any R there is an inverse, written $R^{-1} = R^T$ such that $R\,R^{-1} = R^{-1}\,R = I$.

The rotation group is a subgroup of the more general Galilean group described in Section 9.1.1 of Chapter 9. We will see below that the rotation matrices R are described by three parameters, and so this is a three-parameter group.

There are several ways to describe the relative orientation of two coordinate frames. Some of the common ones are: an axis and angle of rotation, denoted by $(\hat{\mathbf{n}}, \theta)$, Euler angles, denoted by three angles (α, β, γ), and the Cayley-Kline parameters. We will discuss these parameterizations in this section.

In addition, there are two alternative ways to describe rotations: the *active* way, where a point in space is transformed into a new point and which we can think of as a physical rotation of a vector or object, and the *passive* way, where a point remains fixed and the coordinate system is rotated. We use passive rotation here, which was our convention for the general Galilean transformations of Chapter 9. Edmonds [2] uses passive rotation, whereas Biedenharn [5], Rose [6], and Merzbacher [7] all use active rotations.[1]

21.2.1 Rotation matrices

Let Σ and Σ' be two coordinate systems with a common origin, and let a point P described by a vector $\mathbf{r}$ from the origin to the point and let (x, y, z) be Cartesian coordinates of the point in Σ and (x', y', z') be Cartesian coordinates of the *same* point in Σ'. Let us further assume that both of these coordinate systems are oriented in a *right* handed sense.[2] Then we can write the vector $\mathbf{r}$ in either coordinate system using unit vectors:[3]

$$\mathbf{r} = x_i\,\hat{\mathbf{e}}_i = x'_i\,\hat{\mathbf{e}}'_i\,, \tag{21.59}$$

where $\hat{\mathbf{e}}_i$ and $\hat{\mathbf{e}}'_i$ are orthonormal sets of unit vectors describing the two Cartesian coordinate systems: $\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \hat{\mathbf{e}}'_i \cdot \hat{\mathbf{e}}'_j = \delta_{ij}$. So we find that components of the vector $\mathbf{r}$ in the two systems are related by:

$$x'_i = R_{ij}\,x_j\,, \qquad \text{where} \qquad R_{ij} = \hat{\mathbf{e}}'_i \cdot \hat{\mathbf{e}}_j\,, \tag{21.60}$$

where R must satisfy the orthogonal property:

$$R^T_{ik}\,R_{kj} = R_{ki}\,R_{kj} = \delta_{ij}\,. \tag{21.61}$$

That is $R^{-1} = R^T$. The unit vectors transform in the opposite way:

$$\hat{\mathbf{e}}'_i = \hat{\mathbf{e}}_j\,R_{ji} = R^T_{ij}\,\hat{\mathbf{e}}_j\,, \tag{21.62}$$

so that, using the orthogonality relation, Eq. (21.59) is satisfied. From Eq. (21.61) we see that $\det[\,R\,] = \pm 1$, but, in fact, for rotations, we must restrict the determinant to $+1$ since rotations can be generated from the unit matrix, which has a determinant of $+1$.

Matrices describing coordinate systems that are related by *positive* rotations about the x-, y-, and z-axis by an amount α, β, and γ respectively are given by:

$$R_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{pmatrix},\; R_y(\beta) = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix},\; R_z(\gamma) = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{21.63}$$

Notice the location of negative *signs!* One can easily check that these matrices are orthogonal and have determinants of $+1$.

[1] Biedenharn [5] states that the Latin terms for these distinctions are "alibi" for active and "alias" for passive descriptions.
[2] We do not consider space inversions or reflections in this chapter.
[3] In this section, we use a summation convention over repeated indices.

Eq. (21.60) describes a general rotation in terms of nine direction cosines between the coordinate axes,

$$R_{ij} = \hat{\mathbf{e}}'_i \cdot \hat{\mathbf{e}}_j = \cos(\theta_{ij}) \, .$$

These direction cosines, however, are not all independent. The orthogonality requirement, and the fact that the determinant of the matrix must be $+1$, provides six constraint equations, which then leave *three* independent quantities that are needed to describe a rotation.

Exercise 50. Show that if Σ and Σ' are related by a rotation matrix R and Σ' and Σ'' are related by a rotation matrix R', the coordinate systems Σ and Σ'' are related by another orthogonal rotation matrix R''. Find R'' in terms of R and R', and show that it has determinant $+1$.

Definition 33 (The $O^+(3)$ group)**.** The last exercise shows that all three-dimensional rotational matrices R form a three parameter group, called $O^+(3)$, for orthogonal group with positive determinant in three-dimensions.

The direction cosines are not a good way to parameterize the rotation matrices R since there are many relations between the components that are required by orthogonality and unit determinant. In the next sections, we discuss ways to parameterize this matrix.

21.2.2 Axis and angle parameterization

Euler's theorem in classical mechanics states that "the general displacement of a rigid body with one point fixed is a rotation about some axis."[8, p. 156] We show in this section how to parameterize the rotation matrix R by an axis and angle of rotation. We start by writing down the form of the rotation matrix for infinitesimal transformations:

$$R_{ij}(\hat{\mathbf{n}}, \Delta\theta) = \delta_{ij} + \epsilon_{ijk} \hat{n}_k \, \Delta\theta + \cdots \equiv \delta_{ij} + i \, (\, L_k \,)_{ij} \, \hat{n}_k \, \Delta\theta + \cdots \, , \tag{21.64}$$

where $\hat{\mathbf{n}}$ is the axis of rotation, $\Delta\theta$ the magnitude of the rotation. Here we have introduced three imaginary Hermitian and antisymmetric 3×3 matrices $(\, L_k \,)_{ij}$, called the classical generators of the rotation. They are defined by:

$$(\, L_k \,)_{ij} = \frac{1}{i} \, \epsilon_{ijk} \, . \tag{21.65}$$

Explicitly, we have:

$$L_x = \frac{1}{i} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \, , \qquad L_y = \frac{1}{i} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \, , \qquad L_z = \frac{1}{i} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \, . \tag{21.66}$$

Note that these angular momentum matrices are *not* the same as the spin one angular momentum matrices S_i found in Eqs. (21.10), even though they are both 3×3 matrices! The matrices L_k are called the **adjoint** representation of the angular momentum generators. The matrix of unit vectors $\mathbf{L}$ is defined by:

$$\mathbf{L} = L_i \, \hat{\mathbf{e}}_i = \frac{1}{i} \begin{pmatrix} 0 & \hat{\mathbf{e}}_3 & -\hat{\mathbf{e}}_2 \\ -\hat{\mathbf{e}}_3 & 0 & \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 & -\hat{\mathbf{e}}_1 & 0 \end{pmatrix} \, . \tag{21.67}$$

so that we can write, in matrix notation:

$$R(\hat{\mathbf{n}}, \Delta\theta) = 1 + i \, \mathbf{L} \cdot \hat{\mathbf{n}} \, \Delta\theta + \cdots \, . \tag{21.68}$$

So $\mathbf{L}^\dagger = -\mathbf{L}^T = \mathbf{L}$. So $R^T(\hat{\mathbf{n}}, \Delta\theta) = 1 - i \, \mathbf{L} \cdot \hat{\mathbf{n}} \, \Delta\theta + \cdots$. The $\mathbf{L}$ matrix is imaginary, but the $R(\hat{\mathbf{n}}, \Delta\theta)$ matrix is still real. The classical angular momentum generators have *no units* and satisfy the commutation relations:

$$[\, L_i, L_j \,] = i \, \epsilon_{ijk} \, L_k \, , \tag{21.69}$$

which is identical to the ones for the quantum angular momentum operator, except for the fact that in quantum mechanics, the angular momentum operator has units and the commutation relations a factor of $\hbar$. There is no quantum mechanics or $\hbar$ here!

Exercise 51. Carefully explain the differences between the adjoint representation of the angular momentum matrices L_i defined here, and the angular momentum matrices S_i discussed in Section 21.1.1. Can you find a unitary transformation matrix U which relates the S_i set to the L_i set?

We can now construct a finite classical transformation matrix $R(\hat{\mathbf{n}},\theta)$ by compounding N infinitesimal transformation of an amount $\Delta\theta = \theta/N$ about a fixed axis $\hat{\mathbf{n}}$. This gives:

$$R(\hat{\mathbf{n}},\theta) = \lim_{N\to\infty}\left[1 + i\,\frac{\hat{\mathbf{n}}\cdot\mathbf{L}\,\theta}{N}\right]^N = e^{i\,\hat{\mathbf{n}}\cdot\mathbf{L}\,\theta}\,. \tag{21.70}$$

The difficulty here is that the matrix of vectors $\mathbf{L}$ appears in the exponent. We understand how to interpret this by expanding the exponent in a power series. In order to do this, we will need to know the value of powers of the L_i matrices. So we compute:

$$\begin{aligned}
(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)_{ij} &= \frac{1}{i}\,n_k\,\epsilon_{ijk}\,,\\
(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)^2_{ij} &= -n_k\,n_{k'}\,\epsilon_{ilk}\,\epsilon_{ljk'} = n_k\,n_{k'}\,\epsilon_{ikl}\,\epsilon_{ljk'} = n_k\,n_{k'}\,(\,\delta_{ij}\delta_{kk'} - \delta_{ik'}\delta_{kj}\,)\\
&= \delta_{ij} - n_i\,n_j \equiv P_{ij}\\
(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)^3_{ij} &= (\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)^2_{il}\,(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)_{lj} = \frac{1}{i}\,(\,\delta_{il} - n_i\,n_l)\,n_k\,\epsilon_{ljk} = \frac{1}{i}\,(\,n_k\,\epsilon_{ijk} - n_i\,n_l\,n_k\,\epsilon_{ljk}\,)\\
&= \frac{1}{i}\,n_k\,\epsilon_{ijk} = (\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)_{ij}\,,\\
(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)^4_{ij} &= (\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)^2_{ij} = P_{ij}\,, \qquad \text{etc}\cdots
\end{aligned} \tag{21.71}$$

One can see that terms in a power series expansion of $R(\hat{\mathbf{n}},\theta)$ reproduce themselves, so we can collect terms and find:

$$\begin{aligned}
R_{ij}(\hat{\mathbf{n}},\theta) &= \left[\,e^{i\,\theta\hat{\mathbf{n}}\cdot\mathbf{L}}\,\right]_{ij}\\
&= \delta_{ij} + i\,(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)_{ij}\,\theta - \frac{1}{2!}\,(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)^2_{ij}\,\theta^2 - \frac{i}{3!}\,(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)^3_{ij}\,\theta^3 + \frac{1}{4!}\,(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)^4_{ij}\,\theta^4 + \cdots\\
&= n_i\,n_j + P_{ij} + i\,(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)_{ij}\,\theta - \frac{1}{2!}\,P_{ij}\,\theta^2 - \frac{i}{3!}\,(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)_{ij}\,\theta^3 + \frac{1}{4!}\,P_{ij}\,\theta^4 + \cdots\\
&= n_i\,n_j + P_{ij}\,\cos(\theta) + i\,(\,\hat{\mathbf{n}}\cdot\mathbf{L}\,)_{ij}\,\sin(\theta)\\
&= n_i\,n_j + (\,\delta_{ij} - n_i\,n_j\,)\,\cos(\theta) + \epsilon_{ijk}\,n_k\,\sin(\theta)\,.
\end{aligned} \tag{21.72}$$

In terms of unit vectors, the last line can be written as:

$$\begin{aligned}
R_{ij}(\hat{\mathbf{n}},\theta) &= (\,\hat{\mathbf{n}}\cdot\hat{\mathbf{e}}_i\,)\,(\,\hat{\mathbf{n}}\cdot\hat{\mathbf{e}}_j\,) + \left[\,(\,\hat{\mathbf{e}}_i\cdot\hat{\mathbf{e}}_j\,) - (\,\hat{\mathbf{n}}\cdot\hat{\mathbf{e}}_i\,)\,(\,\hat{\mathbf{n}}\cdot\hat{\mathbf{e}}_j\,)\,\right]\cos(\theta) + (\,\hat{\mathbf{n}}\times\hat{\mathbf{e}}_i\,)\cdot\hat{\mathbf{e}}_j\,\sin(\theta)\\
&= (\,\hat{\mathbf{n}}\cdot\hat{\mathbf{e}}_i\,)\,(\,\hat{\mathbf{n}}\cdot\hat{\mathbf{e}}_j\,) + \left[\,(\,\hat{\mathbf{n}}\times(\,\hat{\mathbf{e}}_i\times\hat{\mathbf{n}}\,)\,)\cdot\hat{\mathbf{e}}_j\,\right]\cos(\theta) + (\,\hat{\mathbf{n}}\times\hat{\mathbf{e}}_i\,)\cdot\hat{\mathbf{e}}_j\,\sin(\theta)\,.
\end{aligned} \tag{21.73}$$

So since $\mathbf{r} = x_i\,\hat{\mathbf{e}}_i$, we have:

$$\begin{aligned}
x'_i = R_{ij}(\hat{\mathbf{n}},\theta)\,x_j &= (\,\hat{\mathbf{n}}\cdot\hat{\mathbf{e}}_i\,)\,(\,\hat{\mathbf{n}}\cdot\mathbf{r}\,) + \left[\,(\,\hat{\mathbf{n}}\times(\,\hat{\mathbf{e}}_i\times\hat{\mathbf{n}}\,)\,)\cdot\mathbf{r}\,\right]\cos(\theta) + (\,\hat{\mathbf{n}}\times\hat{\mathbf{e}}_i\,)\cdot\mathbf{r}\,\sin(\theta)\\
&= \left[\,(\,\hat{\mathbf{n}}\cdot\mathbf{r}\,)\,\hat{\mathbf{n}} + (\,\hat{\mathbf{n}}\times(\,\mathbf{r}\times\hat{\mathbf{n}}\,)\,)\,\cos(\theta) + (\,\mathbf{r}\times\hat{\mathbf{n}}\,)\,\sin(\theta)\,\right]\cdot\hat{\mathbf{e}}_i\,,
\end{aligned} \tag{21.74}$$

So if we define $\mathbf{r}'$ as a vector with components in the frame Σ', but with unit vectors in the frame Σ, we find:

$$\mathbf{r}' = x'_i\,\hat{\mathbf{e}}_i = (\,\hat{\mathbf{n}}\cdot\mathbf{r}\,)\,\hat{\mathbf{n}} + (\,\hat{\mathbf{n}}\times(\,\mathbf{r}\times\hat{\mathbf{n}}\,)\,)\,\cos(\theta) + (\,\mathbf{r}\times\hat{\mathbf{n}}\,)\,\sin(\theta)\,. \tag{21.75}$$

Exercise 52. Consider the case of a rotation about the z-axis by an amount θ, so that $\hat{\mathbf{n}} = \hat{\mathbf{e}}_z$, and set $\mathbf{r} = x\,\hat{\mathbf{e}}_x + y\,\hat{\mathbf{e}}_y + z\,\hat{\mathbf{e}}_z$, show that the components of the vector $\mathbf{r}'$, given by Eq. (21.75), are given by $x'_i = R_{ij}(\hat{\mathbf{e}}_z, \theta)\, x_j$, as required.

Exercise 53. Show that the trace of $R(\hat{\mathbf{n}}, \theta)$ gives:

$$\sum_i R_{ii}(\hat{\mathbf{n}}, \theta) = 1 + 2\,\cos(\theta) = 2\,\cos^2(\theta/2)\,, \tag{21.76}$$

where θ is the rotation angle.

Exercise 54. Find the eigenvalues and eigenvectors of $R_{ij}(\hat{\mathbf{e}}_z, \theta)$. Normalize the eigenvectors to the unit sphere, $x^2 + y^2 + z^2 = 1$, and show that the eigenvector with eigenvalue of $+1$ describes the axis of rotation. Extra credit: show that the eigenvalues of an arbitrary orthogonal rotation matrix R are $+1$, 0, and -1. (See Goldstein [8].).

Exercise 55. For the double rotation $R'R = R''$, show that the rotation angle θ'' for the combined rotation is given by:

$$\begin{aligned} 2\,\cos^2(\theta''/2) = (\hat{\mathbf{n}}' \cdot \hat{\mathbf{n}})^2 + 2\,(\hat{\mathbf{n}}' \cdot \hat{\mathbf{n}})^2\,\cos(\theta' + \theta) \\ + \left[\,1 - (\hat{\mathbf{n}}' \cdot \hat{\mathbf{n}})^2\,\right]\left[\,\cos(\theta') + \cos(\theta')\cos(\theta) + \cos(\theta)\,\right]. \end{aligned} \tag{21.77}$$

It is more difficult to find the new axis of rotation $\hat{\mathbf{n}}''$. One way is to find the eigenvector with unit eigenvalue of the resulting matrix, which can be done numerically. There appears to be no closed form for it.

21.2.3 Euler angles

The Euler angles are another way to relate two coordinate systems which are rotated with respect to one another. We define these angles by the following sequence of rotations, which, taken in order, are:[4]

1. Rotate from frame Σ to frame Σ' an angle α about the z-axis, $0 \le \alpha \le 2\pi$.
2. Rotate from frame Σ' to frame Σ'' an angel β about the y'-axis, $0 \le \beta \le \pi$.
3. Rotate from frame Σ'' to frame Σ''' an angle γ about the z''-axis, $0 \le \gamma \le 2\pi$.

The Euler angles are shown in the Fig 21.1. For this definition of the Euler angles, the y'-axis is called the "line of nodes." The coordinates of a *fixed* point P in space, a *passive* rotation, is defined by: (x, y, z) in Σ, (x', y', z') in Σ', (x'', y'', z'') in Σ'', and $(X, Y, Z) \equiv (x''', y''', z''')$ in Σ'''. Then, in a matrix notation,

$$x''' = R_z(\gamma)\,x'' = R_z(\gamma)\,R_y(\beta)\,x' = R_z(\gamma)\,R_y(\beta)\,R_z(\alpha)\,x \equiv R(\gamma, \beta, \alpha)\,x\,, \tag{21.78}$$

where

$$\begin{aligned} R(\gamma, \beta, \alpha) &= R_z(\gamma)\,R_y(\beta)\,R_z(\alpha) \\ &= \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\gamma\cos\beta\cos\alpha - \sin\gamma\sin\alpha, & \cos\gamma\cos\beta\sin\alpha + \sin\gamma\sin\alpha, & -\cos\gamma\sin\beta \\ -\sin\gamma\cos\beta\cos\alpha - \cos\gamma\sin\alpha, & -\sin\gamma\cos\beta\sin\alpha + \cos\gamma\cos\alpha, & \sin\gamma\sin\beta \\ \sin\beta\cos\alpha, & \sin\beta\sin\alpha, & \cos\beta \end{pmatrix}. \end{aligned} \tag{21.79}$$

Here we have used the result in Eqs. (21.63). The rotation matrix $R(\gamma, \beta, \alpha)$ is real, orthogonal, and the determinant is $+1$.

[4] This is the definition of Euler angles used by Edmonds [2][p. 7] and seems to be the most common one for quantum mechanics. In classical mechanics, the second rotation is often about the x'-axis (see Goldstein [8]). Mathematica uses rotations about the x'-axis. Other definitions are often used for the quantum mechanics of a symmetrical top (see Bohr).

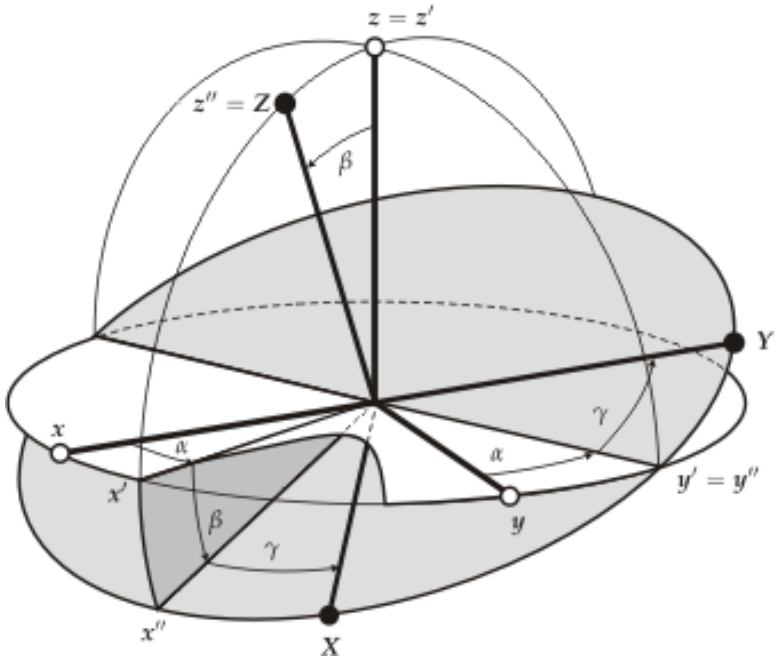

Figure 21.1: Euler angles for the rotations $\Sigma \to \Sigma' \to \Sigma'' \to \Sigma'''$. The final axis is labeled (X, Y, Z).

We will also have occasion to use the inverse of this transformation:

$$\begin{aligned} R^{-1}(\gamma,\beta,\alpha) &= R^T(\gamma,\beta,\alpha) = R_z(-\alpha)\,R_y(-\beta)\,R_z(-\gamma) \\ &= \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma, & -\cos\alpha\cos\beta\sin\gamma + \sin\alpha\sin\gamma, & \cos\alpha\sin\beta \\ \sin\alpha\cos\beta\cos\gamma + \cos\alpha\sin\gamma, & -\sin\alpha\cos\beta\sin\gamma + \cos\alpha\cos\gamma, & \sin\alpha\sin\beta \\ -\sin\beta\cos\gamma, & \sin\beta\sin\gamma, & \cos\beta \end{pmatrix} . \end{aligned} \tag{21.80}$$

We note that the coordinates (x, y, z) in the fixed frame Σ of a point P on the unit circle on z'''-axis in the Σ''' frame, $(x''', y''', z''') = (0, 0, 1)$ is given by:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R^{-1}_{ij}(\alpha,\beta,\gamma) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin\beta\cos\alpha \\ \sin\beta\sin\alpha \\ \cos\beta \end{pmatrix} , \tag{21.81}$$

so the polar angles (θ, ϕ) of this point in the Σ frame is $\theta = \beta$ and $\phi = \alpha$. We will use this result later.

21.2.4 Cayley-Klein parameters

A completely different way to look at rotations is to describe them as directed great circle arcs on the unit sphere in three dimensions. Points on the sphere are described by the set of real variables (x_1, x_2, x_3), with $x_1^2 + x_2^2 + x_3^2 = 1$. These arcs are called **turns** by Biedenharn [5][Ch. 4], and are based on Hamilton's theory of quanterions [9]. Points at the beginning and end of the arc form two reflection planes with the center of the sphere. The line joining these planes is the axis of the rotation and the angle between the planes *half* the angle of rotation. In this section, we adopt the notion of turns as the *active* rotation of vectors, rather

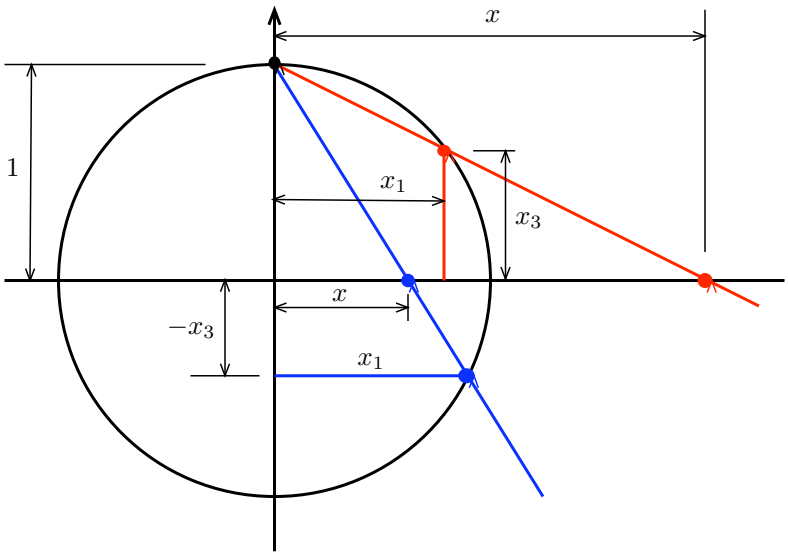

Figure 21.2: Mapping of points on a unit sphere to points on the equatorial plane, for $x_3 > 0$ (red lines) and $x_3 < 0$ (blue lines).

than the passive rotation used in the remainder of this chapter.[5] Turns can be added much like vectors, the geometric rules for which are given by Biedenharn [5][p. 184]. Now a stereographic projection from the North pole of a point on the unit sphere and the equatorial plane maps a unique point on the sphere (except the North pole) to a unique point on the plane, which is described by a *complex* number $z = x + iy$. The geometric mapping can easily be found from Fig. 21.2 by similar triangles to be:

$$z = x + iy = \frac{x_1 + i\,x_2}{1 - x_3} = \frac{1 + x_3}{x_1 - ix_2}\,. \tag{21.82}$$

The upper hemisphere is mapped to points outside the unit circle on the plane, and the lower hemisphere is mapped to points inside the unit circle. Klein [10, 11] and Cayley [12] discovered that a turn, or the rotation of a vector on the unit circle could be described on the plane by a linear fractional transformation of the form:

$$z' = \frac{a\,z + b}{c\,z + d}\,, \tag{21.83}$$

where (a, b, c, d) are complex numbers satisfying:

$$|a|^2 + |b|^2 = |c|^2 + |d|^2 = 1\,, \qquad c\,a^* + d\,b^* = 0\,. \tag{21.84}$$

The set of numbers (a, b, c, d) are called the *Cayley-Klein* parameters. In order to prove this, we need a way to describe turns on the unit sphere. Let $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ be unit vectors describing the start and end point of the turn. Then we can form a scalar $\xi_0 = \hat{\mathbf{r}} \cdot \hat{\mathbf{p}} \equiv \cos(\theta/2)$ and a vector $\boldsymbol{\xi} = \hat{\mathbf{r}} \times \hat{\mathbf{p}} \equiv \hat{\mathbf{n}} \sin(\theta/2)$, which satisfy the property:

$$\xi_0^2 + \xi^2 = 1\,. \tag{21.85}$$

[5]This is the common convention for Cayley-Klein parameters so that the composition rule is satisfied by quaternion multiplication.

Thus a turn can be put in one-to-one correspondence with the set of four quantities $(\xi_0, \boldsymbol{\xi})$ lying on a *four*-dimensional sphere. The rule for addition of a sequence of turns can be found from these definitions. Let $\hat{\mathbf{r}}$, $\hat{\mathbf{p}}$, be unit vectors for the start and end of the first turn described by the parameters $(\xi_0, \boldsymbol{\xi})$, and $\hat{\mathbf{p}}$ and $\hat{\mathbf{s}}$ be the start and end of the second turn described by the parameters $(\xi_0', \boldsymbol{\xi}')$. This means that:

$$\hat{\mathbf{p}} = \xi_0\,\hat{\mathbf{r}} + \boldsymbol{\xi} \times \hat{\mathbf{r}}\,, \qquad \xi_0 = \hat{\mathbf{r}} \cdot \hat{\mathbf{p}}\,, \qquad \boldsymbol{\xi} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}\,, \tag{21.86}$$

$$\hat{\mathbf{s}} = \xi_0'\,\hat{\mathbf{p}} + \boldsymbol{\xi}' \times \hat{\mathbf{p}}\,, \qquad \xi_0' = \hat{\mathbf{p}} \cdot \hat{\mathbf{s}}\,, \qquad \boldsymbol{\xi}' = \hat{\mathbf{p}} \times \hat{\mathbf{s}}\,. \tag{21.87}$$

Substituting (21.86) into (21.87) gives:

$$\begin{aligned} \hat{\mathbf{s}} &= \xi_0'\left(\xi_0\,\hat{\mathbf{r}} + \boldsymbol{\xi} \times \hat{\mathbf{r}}\right) + \boldsymbol{\xi}' \times \left(\xi_0\,\hat{\mathbf{r}} + \boldsymbol{\xi} \times \hat{\mathbf{r}}\right) \\ &= \xi_0''\,\hat{\mathbf{r}} + \boldsymbol{\xi}'' \times \hat{\mathbf{r}}\,, \end{aligned} \tag{21.88}$$

where

$$\begin{aligned} \xi_0'' &= \xi_0'\,\xi_0 - \boldsymbol{\xi}' \cdot \boldsymbol{\xi}\,, \\ \boldsymbol{\xi}'' &= \xi_0\,\boldsymbol{\xi}' + \xi_0'\,\boldsymbol{\xi} + \boldsymbol{\xi}' \times \boldsymbol{\xi}\,. \end{aligned} \tag{21.89}$$

Now since $\hat{\mathbf{r}} \cdot \boldsymbol{\xi}'' = 0$, we find from (21.88) that

$$\hat{\mathbf{r}} \cdot \hat{\mathbf{s}} = \cos(\theta''/2)\,, \qquad \hat{\mathbf{r}} \times \hat{\mathbf{s}} = \hat{\mathbf{n}}''\,\sin(\theta''/2)\,, \tag{21.90}$$

which means that the set of all turns form a group, with a composition rule.

Exercise 56. Show that (21.89) follows from (21.88). Show also that $\hat{\mathbf{r}} \cdot \boldsymbol{\xi}'' = 0$.

Cayley [13] noticed that the composition rule, Eq. (21.89), is the same rule for as the rule for multiplication of two quaternions. That is, if we define

$$\hat{\xi} = \xi_0\,\hat{1} + \xi_1\,\hat{\imath} + \xi_2\,\hat{\jmath} + \xi_3\,\hat{k} = \xi_0\,\hat{1} + \boldsymbol{\xi}\,, \tag{21.91}$$

where the quaternion multiplication rules are:[6]

$$\hat{\imath}\,\hat{\jmath} = -\hat{\jmath}\,\hat{\imath} = \hat{k}\,, \qquad \hat{\jmath}\,\hat{k} = -\hat{k}\,\hat{\jmath} = \hat{\imath}\,, \qquad \hat{k}\,\hat{\imath} = -\hat{\imath}\,\hat{k} = \hat{\jmath}\,, \qquad \hat{1}^2 = \hat{1}\,, \qquad \hat{\imath}^2 = \hat{\jmath}^2 = \hat{k}^2 = -\hat{1}\,, \tag{21.92}$$

then it is easy to show that *quaternion* multiplication:

$$\hat{\xi}'' = \hat{\xi}'\,\hat{\xi}\,, \tag{21.93}$$

reproduces the composition rule (21.89). So it is natural to use the algebra of quaternions to describe rotations.

Exercise 57. Show that Eq. (21.93) reproduces the composition rule (21.89) using the quaternion multiplication rules of Eq. (21.92).

Definition 34 (adjoint quaternion)**.** The adjoint quaternion $\hat{\xi}^\dagger$ is defined by:

$$\hat{\xi}^\dagger = \xi_0\,\hat{1} - \xi_1\,\hat{\imath} - \xi_2\,\hat{\jmath} - \xi_3\,\hat{k} = \xi_0\,\hat{1} - \boldsymbol{\xi}\,, \tag{21.94}$$

so that the length of $\hat{\xi}$ is given by:

$$\hat{\xi}^\dagger\,\hat{\xi} = \xi_0^2 + \xi^2 = 1\,. \tag{21.95}$$

[6] One should think of quaternions as an extension of the complex numbers. They form what is called a **division algebra**. We designate quaternions with a hat symbol.

Now let $\hat{r}$ be a quaternion describing the position of a vector (x_1, x_2, x_3) on the unit sphere, defined by:

$$\hat{r} = x_1\,\hat{i} + x_2\,\hat{j} + x_3\,\hat{k} = \hat{\mathbf{r}}\,, \qquad \text{with} \qquad x_0 = 0\,, \qquad x_1^2 + x_2^2 + x_3^2 = 1\,. \tag{21.96}$$

Then the rotation of the vector $\hat{\mathbf{r}}$ on the unit sphere is described by the quaternion product: $\hat{r}' = \hat{\xi}\,\hat{r}\,\hat{\xi}^\dagger$. We state this in the next theorem:

Theorem 41 (quaternion rotations)**.** *The rotation of a vector $\hat{\mathbf{r}}$ on the unit sphere is given by the quaternion product:*

$$\hat{r}' = \hat{\xi}\,\hat{r}\,\hat{\xi}^\dagger\,. \tag{21.97}$$

Proof. From (21.97) and the composition rules (21.89), we find:

$$\begin{aligned}
\hat{r}' &= (\,\xi_0 + \boldsymbol{\xi}\,)\,\hat{\mathbf{r}}\,(\,\xi_0 - \boldsymbol{\xi}\,) \\
&= (\,\xi_0 + \boldsymbol{\xi}\,)\,(\,\hat{\mathbf{r}}\cdot\boldsymbol{\xi} + \xi_0\,\hat{\mathbf{r}} - \hat{\mathbf{r}}\times\boldsymbol{\xi}\,) \\
&= \xi_0\,(\,\hat{\mathbf{r}}\cdot\boldsymbol{\xi}\,) - \boldsymbol{\xi}\cdot(\,\xi_0\,\hat{\mathbf{r}} - \hat{\mathbf{r}}\times\boldsymbol{\xi}\,) + \xi_0\,(\,\xi_0\,\hat{\mathbf{r}} - \hat{\mathbf{r}}\times\boldsymbol{\xi}\,) + (\,\hat{\mathbf{r}}\cdot\boldsymbol{\xi}\,)\,\boldsymbol{\xi} + \boldsymbol{\xi}\times(\,\xi_0\,\hat{\mathbf{r}} - \hat{\mathbf{r}}\times\boldsymbol{\xi}\,) \\
&= (\,\xi_0^2 - \xi^2\,)\,\hat{\mathbf{r}} - 2\,\xi_0\,(\,\hat{\mathbf{r}}\times\boldsymbol{\xi}\,) + 2\,(\,\hat{\mathbf{r}}\cdot\boldsymbol{\xi}\,)\,\boldsymbol{\xi} \\
&= \hat{\mathbf{r}}\,\cos(\theta) - (\,\hat{\mathbf{r}}\times\hat{\mathbf{n}}\,)\,\sin(\theta) + (\,\hat{\mathbf{r}}\cdot\hat{\mathbf{n}}\,)\,\hat{\mathbf{n}}\,(\,1 - \cos(\theta)\,) \\
&= (\,\hat{\mathbf{r}}\cdot\hat{\mathbf{n}}\,)\,\hat{\mathbf{n}} + (\,\hat{\mathbf{r}} - (\,\hat{\mathbf{r}}\cdot\hat{\mathbf{n}}\,)\,\hat{\mathbf{n}}\,)\,\cos(\theta) - (\,\hat{\mathbf{r}}\times\hat{\mathbf{n}}\,)\,\sin(\theta) \\
&= (\,\hat{\mathbf{r}}\cdot\hat{\mathbf{n}}\,)\,\hat{\mathbf{n}} + (\,\hat{\mathbf{n}}\times(\,\hat{\mathbf{r}}\times\hat{\mathbf{n}}\,)\,)\,\cos(\theta) - (\,\hat{\mathbf{r}}\times\hat{\mathbf{n}}\,)\,\sin(\theta) = \hat{\mathbf{r}}'\,,
\end{aligned} \tag{21.98}$$

where $\hat{\mathbf{r}}' = x_1'\,\hat{i} + x_2'\,\hat{j} + x_3'\,\hat{k}$. From Eq. (21.75), we recognize $\hat{\mathbf{r}}'$, as the *active* rotation of a vector $\hat{\mathbf{r}}$ in a fixed coordinate system about an axis $\hat{\mathbf{n}}$ by an amount θ, which is what we were trying to prove. □

Rather than using quaternions, physicists often prefer to use the Pauli matrices to represent turns. That is, if we introduce the mapping,

$$\hat{1} \mapsto 1\,, \qquad \hat{i} \mapsto -i\sigma_x\,, \qquad \hat{j} \mapsto -i\sigma_y\,, \qquad \hat{k} \mapsto -i\sigma_z\,, \tag{21.99}$$

so that a turn $\hat{\xi}$ is represented by a *unitary* 2×2 matrix:[7]

$$\hat{\xi} \mapsto D(\xi) = \xi_0 - i\,\boldsymbol{\xi}\cdot\boldsymbol{\sigma} = \cos(\theta/2) - i\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,\sin(\theta/2) = e^{-i\,\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\,\theta/2}\,. \tag{21.100}$$

Here the quaternion composition rule is represented by *matrix* multiplication. The factor $-i$ in the mapping (21.99) is necessary to get the correct composition rule. We prove this in Exercise 58 below. Since the Pauli matrices are Hermitian, $\hat{\xi}^\dagger \mapsto D(\xi^\dagger) = \xi_0 + i\,\boldsymbol{\xi}\cdot\boldsymbol{\sigma} = D^\dagger(\xi)$, as expected.

Exercise 58. Prove that matrix multiplication $D(\xi'') = D(\xi')\,D(\xi)$ yields the same composition rule as the quaternion composition rule given in Eq. (21.89).

A point P on the unit sphere is now represented by a 2×2 matrix function of coordinates given by:

$$\hat{r} \mapsto -i\,\hat{\mathbf{r}}\cdot\boldsymbol{\sigma} = -i\begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix}\,. \tag{21.101}$$

with a similar expression for the rotated vector $\hat{r}' \mapsto -i\,\hat{\mathbf{r}}'\cdot\boldsymbol{\sigma}$ on the unit sphere in a fixed coordinate system. Then the matrix version of the quaternion rotation of Theorem 41 is given in the next theorem.

Theorem 42. *The rotation of a vector $\hat{\mathbf{r}}$ on the unit sphere is given by the matrix product:*

$$\hat{\mathbf{r}}'\cdot\boldsymbol{\sigma} = D(\xi)\,\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}\,D^\dagger(\xi)\,, \tag{21.102}$$

where $D(\xi)$ is given by Eq. (21.100).

[7] The $D(\xi)$ matrix defined in this section is for *active* rotations.

Proof. Using the properties of the Pauli matrices, we first work out:

$$\begin{aligned} D(\xi)\,\boldsymbol{\sigma}\,D^{\dagger}(\xi) &= \big[\cos(\theta/2) - i\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,\sin(\theta/2)\big]\,\boldsymbol{\sigma}\,\big[\cos(\theta/2) + i\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,\sin(\theta/2)\big] \\ &= \boldsymbol{\sigma}\,\cos^2(\theta/2) + i\,[\boldsymbol{\sigma},(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})]\,\sin(\theta/2)\,\cos(\theta/2) + (\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,\boldsymbol{\sigma}\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,\sin^2(\theta/2)\,. \end{aligned} \tag{21.103}$$

Using:

$$\begin{aligned} [\boldsymbol{\sigma},(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})] &= 2i\,(\hat{\mathbf{n}}\times\boldsymbol{\sigma})\,, \\ (\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,\boldsymbol{\sigma}\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}) &= \boldsymbol{\sigma} + 2i\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,(\hat{\mathbf{n}}\times\boldsymbol{\sigma}) = 2\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,\hat{\mathbf{n}} - \boldsymbol{\sigma}\,, \end{aligned} \tag{21.104}$$

then Eq. (21.103) becomes:

$$\begin{aligned} D(\xi)\,\boldsymbol{\sigma}\,D^{\dagger}(\xi) &= \boldsymbol{\sigma}\,\cos(\theta) - (\,\hat{\mathbf{n}}\times\boldsymbol{\sigma}\,)\,\sin(\theta) + (\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,\hat{\mathbf{n}}\,(1-\cos(\theta)\,) \\ &= (\,\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\,)\,\hat{\mathbf{n}} + \hat{\mathbf{n}}\times(\,\boldsymbol{\sigma}\times\hat{\mathbf{n}}\,)\,\cos(\theta) - (\,\hat{\mathbf{n}}\times\boldsymbol{\sigma}\,)\,\sin(\theta)\,. \end{aligned} \tag{21.105}$$

So (21.102) is given by:

$$\begin{aligned} \hat{\mathbf{r}}'\cdot\boldsymbol{\sigma} &= D(\xi)\,\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}\,D^{\dagger}(\xi) \\ &= (\,\hat{\mathbf{n}}\cdot\hat{\mathbf{r}}\,)\,(\,\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\,) + (\,\hat{\mathbf{r}}\cdot\boldsymbol{\sigma} - (\hat{\mathbf{r}}\cdot\hat{\mathbf{n}})\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,)\,\cos(\theta) - \hat{\mathbf{r}}\cdot(\,\hat{\mathbf{n}}\times\boldsymbol{\sigma}\,)\,\sin(\theta) \\ &= \big[\,(\,\hat{\mathbf{n}}\cdot\hat{\mathbf{r}}\,)\,\hat{\mathbf{n}} + \hat{\mathbf{n}}\times(\,\hat{\mathbf{r}}\times\hat{\mathbf{n}}\,)\,\cos(\theta) - (\,\hat{\mathbf{r}}\times\hat{\mathbf{n}}\,)\,\sin(\theta)\big]\cdot\boldsymbol{\sigma} = \hat{\mathbf{r}}'\cdot\boldsymbol{\sigma}\,, \end{aligned} \tag{21.106}$$

where $\hat{\mathbf{r}}'$ is given by (21.98). This completes the proof. □

Exercise 59. Show that $\det[\,\hat{\mathbf{r}}'\cdot\boldsymbol{\sigma}\,] = \det[\,\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}\,] = 1$.

But Theorem 42 is not the only way to describe the rotation of a vector. We can also use the transformation properties of spinors which are eigenvectors of the operator $\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}$. This is the content of the next theorem.

Theorem 43 (Cayley-Klein rotation)**.** *The rotation of a vector on the unit sphere can be described by a linear fractional transformation on the plane of the form:*

$$z' = \frac{a\,z+b}{c\,z+d}\,, \tag{21.107}$$

where z is given by:

$$z = x + iy = \frac{x_1 + i\,x_2}{1-x_3}\,, \tag{21.108}$$

and where (a,b,c,d) satisfy:

$$|a|^2 + |b|^2 = |c|^2 + |d|^2 = 1\,, \qquad c\,a^* + d\,b^* = 0\,. \tag{21.109}$$

Proof. From the results of Theorem 42, Eq. (21.102) gives:

$$\hat{\mathbf{r}}'\cdot\boldsymbol{\sigma}\,D(\xi) = D(\xi)\,\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}\,, \tag{21.110}$$

since $D(\xi)$ is unitary. Now the matrix $\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}$ is Hermitian and has two eigenvalues and eigenvectors. That is:

$$\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}\,\chi_\lambda(\hat{\mathbf{r}}) = \lambda\,\chi_\lambda(\hat{\mathbf{r}})\,, \qquad \text{with} \qquad \lambda = \pm 1\,. \tag{21.111}$$

One can easily check that the eigenvectors are given by (see Section 15.2.1):

$$\chi_+(\hat{\mathbf{r}}) = \mathcal{N}_+\begin{pmatrix} x_1 - i x_2 \\ 1 - x_3 \end{pmatrix}\,, \qquad \text{and} \qquad \chi_-(\hat{\mathbf{r}}) = \mathcal{N}_-\begin{pmatrix} x_3 - 1 \\ x_1 + i x_2 \end{pmatrix}\,, \tag{21.112}$$

where $\mathcal{N}_\pm$ are normalization factors. So, from (21.110), we find:

$$\hat{\mathbf{r}}' \cdot \boldsymbol{\sigma} \left\{ D(\xi)\, \chi_\lambda(\hat{\mathbf{r}}) \right\} = \lambda \left\{ D(\xi)\, \chi_\lambda(\hat{\mathbf{r}}) \right\} , \tag{21.113}$$

from which we conclude that:

$$\chi_\lambda(\hat{\mathbf{r}}') = \mathcal{N}\, D(\xi)\, \chi_\lambda(\hat{\mathbf{r}}) , \tag{21.114}$$

where $\mathcal{N}$ is some constant. Now from (21.100),

$$D(\xi) = \xi_0 - i\, \boldsymbol{\xi} \cdot \boldsymbol{\sigma} = \begin{pmatrix} \xi_0 - i\xi_3 & -i\xi_1 + \xi_2 \\ -i\xi_1 - \xi_2 & \xi_0 + i\xi_3 \end{pmatrix} \equiv \begin{pmatrix} a^* & b^* \\ c^* & d^* \end{pmatrix} \tag{21.115}$$

which defines the parameters (a, b, c, d). From the unitarity of the $D(\xi)$ matrix, we easily establish that:

$$|a|^2 + |b|^2 = |c|^2 + |d|^2 = 1 , \qquad c\, a^* + d\, b^* = 0 . \tag{21.116}$$

So then for the $\lambda = +1$ eigenvector, Eq. (21.114) gives:

$$\begin{aligned} \mathcal{N}'_+ \left(x'_1 - i x'_2 \right) &= \mathcal{N}\, \mathcal{N}_+ \left[a^* \left(x_1 - i x_2 \right) + b^* \left(1 - x_3 \right) \right] \\ \mathcal{N}'_+ \left(1 - x'_3 \right) &= \mathcal{N}\, \mathcal{N}_+ \left[c^* \left(x_1 - i x_2 \right) + d^* \left(1 - x_3 \right) \right] . \end{aligned} \tag{21.117}$$

The complex conjugate of the ratio of the first and second equations of (21.117) gives:

$$z' = \frac{a\, z + b}{c\, z + d} , \qquad \text{where} \qquad z = \frac{x_1 + i x_2}{1 - x_3} , \tag{21.118}$$

which is the result we were trying to prove. We leave investigation of the $\lambda = -1$ eigenvector to Exercise 60. □

Exercise 60. Show that the $\lambda = -1$ eigenvalue given in Eq. (21.112) gives a rotation on the unit sphere for the mapping $z \mapsto -1/z^*$ of the complex plane. This mapping corresponds to a stereographic projection from the *South* pole followed by a negative complex conjugation, which is an equivalent one-to-one mapping of points on the unit sphere to the complex plane.

Theorem 43 establishes the claim by Klein and Cayley that the fractional linear transformation of the complex projection plane, Eq. (21.83), represents a rotation on the unit sphere.

Remark 33. We have exhibited in this section a direct connection between different ways to describe the rotation of vectors. We can use either the rotation matrices R_{ij}, quaternions ξ and the Cayley-Klein parameters (a, b, c, d), and two-dimensional Pauli matrices $D(\xi)$. All of these methods are strictly classical, and provide equivalent means of describing rotated coordinate systems. It should not be surprising that there is this connection, since the 3×3 rotation matrices R belong to the group $O^+(3)$ and the 2×2 unitary matrices $D(\xi)$ belong to the group $SU(2)$. It is well known that these two groups are isomorphic: $O^+(3) \sim SU(2)$. In this section, we have shown how to describe rotations with either representation and the connection between them. We emphasize again that our discussion is completely classical.

Remark 34. Since in the rest of this chapter, we use the *passive* rotation convention where the point in space remains fixed but the coordinate system is rotated, let us write down the Cayley-Klein transformations for passive rotations. We have:

$$\mathbf{r}' \cdot \boldsymbol{\sigma} = D(R)\, \mathbf{r} \cdot \boldsymbol{\sigma}\, D^\dagger(R) , \tag{21.119}$$

where now $\mathbf{r} = x_i\, \hat{\mathbf{e}}_i$ and $\mathbf{r}' = x'_i\, \hat{\mathbf{e}}_i$, with $x'_i = R_{ij}\, x_j$. The $D(R)$ rotation matrix acts only on $\boldsymbol{\sigma}$ and can be written in several ways. In terms of an axis and angle of rotation $(\hat{\mathbf{n}}, \theta)$, the rotation matrix $D(R)$ is given by:

$$D(\hat{\mathbf{n}}, \theta) = e^{i\, \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}\, \theta/2} = \cos(\theta/2) + i\, (\hat{\mathbf{n}} \cdot \boldsymbol{\sigma})\, \sin(\theta/2) , \tag{21.120}$$

in terms of the quaternion $\xi = (\xi_0, \boldsymbol{\xi})$ and the Cayley-Klein parameters (a, b, c, d), it is:

$$D(\xi) = \xi_0 + i\,\boldsymbol{\xi}\cdot\boldsymbol{\sigma} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\,, \tag{21.121}$$

and in terms of the Euler angles (α, β, γ),

$$\begin{aligned} D(\gamma,\beta,\alpha) &= D(\hat{\mathbf{e}}_z,\gamma)\,D(\hat{\mathbf{e}}_y,\beta)\,D(\hat{\mathbf{e}}_z,\alpha) = e^{i\sigma_z\gamma/2}\,e^{i\sigma_y\beta/2}\,e^{i\sigma_z\alpha/2} \\ &= \begin{pmatrix} e^{i(+\gamma+\alpha)/2}\cos(\beta/2) & e^{i(+\gamma-\alpha)/2}\sin(\beta/2) \\ -e^{i(-\gamma+\alpha)/2}\sin(\beta/2) & e^{i(-\gamma-\alpha)/2}\cos(\beta/2) \end{pmatrix}\,, \end{aligned} \tag{21.122}$$

We shall have occasion to use all these different forms.

Remark 35. A tensor T, defined by the expansion $\mathsf{T} = T_{ij}\,\sigma_i\,\sigma_j$, transforms under rotations of the coordinate systems as:

$$\mathsf{T}' = D(R)\,\mathsf{T}\,D^\dagger(R)\,, \tag{21.123}$$

where $\mathsf{T}' = T'_{ij}\,\sigma_i\,\sigma_j$, with $T'_{i,j} = R_{i,i'}\,R_{j,j'}\,T_{i',j'}$. We can generalize this result to tensors of any rank.

21.3 Rotations in quantum mechanics

In quantum mechanics, symmetry transformations, such as rotations of the coordinate system, are represented by unitary transformations of vectors in the Hilbert space. Unitary representations of the rotation group are faithful representations. This means that the composition rule, $R'' = R'\,R$ of the group is preserved by the unitary representation, *without any phase factors.*[8] That is: $U(R'') = U(R')\,U(R)$. We also have $U(1) = 1$ and $U^{-1}(R) = U^\dagger(R) = U(R^{-1})$. For infinitesimal rotations, we write the classical rotational matrix as in Eq. (21.68):

$$R_{ij}(\hat{\mathbf{n}},\Delta\theta) = \delta_{ij} + \epsilon_{ijk}\hat{n}_k\,\Delta\theta + \cdots\,, \tag{21.124}$$

which we abbreviate as $R = 1 + \Delta\theta + \cdots$. We write the infinitesimal unitary transformtion as:

$$U_{\mathbf{J}}(1+\Delta\theta) = 1 + i\,n_i J_i\,\Delta\theta/\hbar + \cdots\,, \tag{21.125}$$

where J_i is the Hermitian generator of the transformation. We will show in this section that the set of generators J_i, for $i = 1, 2, 3$, transform under rotations in quantum mechanics as a pseudo-vector and that it obeys the commutation relations we assumed in Eq. (21.1) at the beginning of this chapter. The factor of $\hbar$ is inserted here so that J_i can have units of classical angular momentum, and is the *only* way that makes $U_{\mathbf{J}}(R)$ into a quantum operator. Now let us consider the combined transformation:

$$U^\dagger_{\mathbf{J}}(R)\,U_{\mathbf{J}}(1+\Delta\theta')\,U_{\mathbf{J}}(R) = U_{\mathbf{J}}(R^{-1})\,U_{\mathbf{J}}(1+\Delta\theta')\,U_{\mathbf{J}}(R) = U_{\mathbf{J}}(\,R^{-1}\,(\,1+\Delta\theta'\,)\,R\,) = U_{\mathbf{J}}(1+\Delta\theta'')\,. \tag{21.126}$$

We first work out the classical transformation:

$$1 + \Delta\theta'' + \cdots = R^{-1}\,(\,1+\Delta\theta'\,)\,R = 1 + R^{-1}\,\Delta\theta'\,R + \cdots \tag{21.127}$$

That is

$$\epsilon_{ijk}\hat{n}_k\,\Delta\theta'' = \epsilon_{i'j'k'}\,R_{i'i}\,R_{j'j}\,\hat{n}_{k'}\,\Delta\theta'\,. \tag{21.128}$$

Now using the relation:

$$\det[\,R\,]\,\epsilon_{ijk} = \epsilon_{i'j'k'}\,R_{i'i}\,R_{j'j}\,R_{k'k}\,, \qquad \text{or} \qquad \det[\,R\,]\,\epsilon_{ijk}\,R_{k'k} = \epsilon_{i'j'k'}\,R_{i'i}\,R_{j'j}\,. \tag{21.129}$$

[8] This is not the case for the full Galilean group, where there is a phase factor involved (see Chapter 9 and particularly Section 9.5).

Inserting this result into (21.128) gives the relation:

$$\hat{n}_k\,\Delta\theta'' = \det[\,R\,]\,R_{k'k}\,\hat{n}_{k'}\,\Delta\theta' \tag{21.130}$$

So from (21.126), we find:

$$\begin{aligned} 1 + i\,\hat{n}_j J_j\,\Delta\theta''/\hbar + \cdots &= U^\dagger_{\mathbf{J}}(R)\left\{\,1 + i\,\hat{n}_i J_i\,\Delta\theta'/\hbar + \cdots\right\} U_{\mathbf{J}}(R) \\ &= 1 + i\,U^\dagger_{\mathbf{J}}(R)\,J_i\,U_{\mathbf{J}}(R)\,\hat{n}_i\,\Delta\theta'/\hbar + \cdots\,, \end{aligned} \tag{21.131}$$

or

$$U^\dagger_{\mathbf{J}}(R)\,J_i\,U_{\mathbf{J}}(R)\,\hat{n}_i\Delta\theta' = \hat{n}_j J_j\,\Delta\theta'' = \det[\,R\,]\,R_{ij}\,J_j\,\hat{n}_i\,\Delta\theta'\,. \tag{21.132}$$

Comparing coefficients of $\hat{n}_i\,\Delta\theta'$ on both sides of this equation, we find:

$$U^\dagger_{\mathbf{J}}(R)\,J_i\,U_{\mathbf{J}}(R) = \det[\,R\,]\,R_{ij}\,J_j\,, \tag{21.133}$$

showing that under rotations, the generators of rotations J_i transform as pseudo-vectors. For ordinary rotations $\det[\,R\,] = +1$; whereas for Parity or mirror inversions of the coordinate system $\det[\,R\,] = -1$. We restrict ourselves here to ordinary rotations. Iterating the infinitesimal rotation operator (21.127) gives the finite unitary transformation:

$$U_{\mathbf{J}}(\hat{\mathbf{n}},\theta) = e^{i\,\hat{\mathbf{n}}\cdot\mathbf{J}\,\theta/\hbar}\,, \qquad R \mapsto (\hat{\mathbf{n}},\theta)\,. \tag{21.134}$$

Further expansion of $U(R)$ in Eq. (21.133) for infinitesimal $R = 1 + \Delta\theta + \cdots$ gives:

$$\left\{\,1 - i\,\hat{n}_j J_j\,\Delta\theta/\hbar + \cdots\right\} J_i \left\{\,1 + i\,\hat{n}_j J_j\,\Delta\theta/\hbar + \cdots\right\} = \left\{\,\delta_{ij} + \epsilon_{ijk}\hat{n}_k\,\Delta\theta + \cdots\right\} J_j\,. \tag{21.135}$$

Comparing coefficients of $\hat{n}_j\,\Delta\theta$ on both sides of this equation gives the commutation relations for the angular momentum generators:

$$[\,J_i, J_j\,] = i\hbar\,\epsilon_{ijk}\,J_k\,. \tag{21.136}$$

This derivation of the properties of the unitary transformations and generators of the rotation group parallels that of the properties of the full Galilean group done in Chapter 9.

Remark 36. When $j = 1/2$ we can put $\mathbf{J} = \mathbf{S} = \hbar\boldsymbol{\sigma}/2$, so that the unitary rotation operator is given by:

$$U_{\mathbf{S}}(\hat{\mathbf{n}},\theta) = e^{i\hat{\mathbf{n}}\cdot\mathbf{S}/\hbar} = e^{i\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}/2}\,, \tag{21.137}$$

which is the same as the unitary operator, Eq. (G.131), which we used to describe classical rotations in the adjoint representation.

Exercise 61. Suppose the composition rule for the unitary representation of the rotation group is of the form:

$$U(R')\,U(R) = e^{i\phi(R',R)}\,U(R'R)\,, \tag{21.138}$$

where $\phi(R',R)$ is a phase which may depend on R and R'. Using Bargmann's method (see Section 9.2.1), show that the phase $\phi(R',R)$ is a trivial phase, and can be absorbed into the overall phase of the unitary transformation. This exercise shows that the unitary representation of the rotation group is faithful.

Now we want to find relations between eigenvectors $|\,j,m\,\rangle$ angular momentum in two frames related by a rotation. So let $|\,j,m\,\rangle$ be eigenvectors of J^2 and J_z in the Σ frame and $|\,j,m\,\rangle'$ be eigenvectors of J^2 and J_z in the Σ' frame. We first note that the square of the total angular momentum vector is invariant under rotations:

$$U^\dagger_{\mathbf{J}}(R)\,J^2\,U_{\mathbf{J}}(R) = J^2\,, \tag{21.139}$$

so the total angular momentum quantum numbers for the eigenvectors must be the same in each frame, $j' = j$. From (21.133), J_i transforms as follows (in the following, we consider the case when $\det[\,R\,] = +1$):

$$U^\dagger_{\mathbf{J}}(R)\,J_i\,U_{\mathbf{J}}(R) = R_{i,j}\,J_j = J'_i\,, \tag{21.140}$$

So multiplying (21.140) on the left by $U^\dagger_{\mathbf{J}}(R)$, setting $i = z$, and operating on the eigenvector $|\, j, m \,\rangle$ defined in frame Σ, we find:

$$J_{z'} \left\{ U^\dagger_{\mathbf{J}}(R) \,|\, j, m \,\rangle \right\} = U^\dagger_{\mathbf{J}}(R)\, J_z \,|\, j, m \,\rangle = \hbar\, m \left\{ U^\dagger_{\mathbf{J}}(R) \,|\, j, m \,\rangle \right\}, \tag{21.141}$$

from which we conclude that $U^\dagger_{\mathbf{J}}(R) \,|\, j, m \,\rangle$ is an eigenvector of $J_{z'}$ with eigenvalue $\hbar\, m$. That is:

$$|\, j, m \,\rangle' = U^\dagger_{\mathbf{J}}(R) \,|\, j, m \,\rangle = \sum_{m'=-j}^{+j} |\, j, m' \,\rangle \,\langle\, j, m' \,|\, U^\dagger_{\mathbf{J}}(R) \,|\, j, m \,\rangle = \sum_{m'=-j}^{+j} D^{(j)\,*}_{m,m'}(R) \,|\, j, m' \,\rangle\,, \tag{21.142}$$

where we have defined the D-functions, which are angular momentum matrix elements of the rotation operator, by:

Definition 35 (D-functions)**.** The D-functions are the matrix elements of the rotation operator, and are defined by:

$$D^{(j)}_{m,m'}(R) = \langle\, j, m \,|\, U_{\mathbf{J}}(R) \,|\, j, m' \,\rangle = {}'\langle\, j, m \,|\, j, m' \,\rangle = {}'\langle\, j, m \,|\, U_{\mathbf{J}}(R) \,|\, j, m' \,\rangle'\,. \tag{21.143}$$

The D-function can be computed in either the Σ or Σ' frames. Eq. (21.142) relates eigenvectors of the angular momentum in frame Σ' to those in Σ. Note that the matrix $D^{(j)}_{m,m'}(R)$ is the overlap between the state $|\, j, m \,\rangle'$ in the Σ' frame and $|\, j, m \,\rangle$ in the Σ frame. The *row's* of this matrix are the adjoint eigenvectors of J'_z in the Σ frame, so that the *columns* of the adjoint matrix, $D^{(j)\,*}_{m',m}(R)$ are the eigenvectors of J'_z in the Σ frame.

For infinitesimal rotations, the D-function is given by:

$$\begin{aligned} D^{(j)}_{m,m'}(\hat{\mathbf{n}}, \Delta\theta) &= \langle\, j, m \,|\, U_{\mathbf{J}}(\hat{\mathbf{n}}, \Delta\theta) \,|\, j, m' \,\rangle = \langle\, j, m \,|\, \Big\{ 1 + \frac{i}{\hbar}\, \hat{\mathbf{n}} \cdot \mathbf{J}\, \Delta\theta + \cdots \Big\} \,|\, j, m' \,\rangle \\ &= \delta_{m,m'} + \frac{i}{\hbar} \,\langle\, j, m \,|\, \hat{\mathbf{n}} \cdot \mathbf{J} \,|\, j, m' \,\rangle\, \Delta\theta + \cdots \end{aligned} \tag{21.144}$$

Exercise 62. Find the first order matrix elements of $D^{(j)}_{m,m'}(\hat{\mathbf{n}}, \Delta\theta)$ for $\hat{\mathbf{n}} = \hat{\mathbf{e}}_z$ and $\hat{\mathbf{n}} = \hat{\mathbf{e}}_x \pm i\hat{\mathbf{e}}_y$.

21.3.1 Rotations using Euler angles

Consider the sequential rotations $\Sigma \to \Sigma' \to \Sigma'' \to \Sigma'''$, described by the Euler angles defined in Section 21.2.3. The unitary operator in quantum mechanics for this classical transformation is then given by the composition rule:

$$U_{\mathbf{J}}(\gamma, \beta, \alpha) = U_{\mathbf{J}}(\hat{\mathbf{e}}_z, \gamma)\, U_{\mathbf{J}}(\hat{\mathbf{e}}_y, \beta)\, U_{\mathbf{J}}(\hat{\mathbf{e}}_z, \alpha) = e^{iJ_z\gamma/\hbar}\, e^{iJ_y\beta/\hbar}\, e^{iJ_z\alpha/\hbar}\,. \tag{21.145}$$

So the angular momentum operator J_i transforms according to ($\det[\, R\,] = 1$):

$$U^\dagger_{\mathbf{J}}(\gamma, \beta, \alpha)\, J_i\, U_{\mathbf{J}}(\gamma, \beta, \alpha) = R_{z\,ij}(\gamma)\, R_{y\,jk}(\beta)\, R_{z\,kl}(\alpha)\, J_l = R_{il}(\gamma, \beta, \alpha)\, J_l \equiv J'''_i\,, \tag{21.146}$$

where $R_{il}(\gamma, \beta, \alpha)$ is given by Eq. (21.79). Again, multiplying on the right by $U^\dagger_{\mathbf{J}}(\gamma, \beta, \alpha)$, setting $i = z$, and operating on the eigenvector $|\, j, m \,\rangle$ defined in frame Σ, we find:

$$J_{z'''} \left\{ U^\dagger_{\mathbf{J}}(\gamma, \beta, \alpha) \,|\, j, m \,\rangle \right\} = U^\dagger_{\mathbf{J}}(\gamma, \beta, \alpha)\, J_z \,|\, j, m \,\rangle = \hbar\, m \left\{ U^\dagger_{\mathbf{J}}(\gamma, \beta, \alpha) \,|\, j, m \,\rangle \right\}. \tag{21.147}$$

So we conclude here that $U^\dagger_{\mathbf{J}}(\alpha, \beta, \gamma) \,|\, j, m \,\rangle$ is an eigenvector of $J_{z'''}$ with eigenvalue $\hbar\, m$. That is:

$$\begin{aligned} |\, j, m \,\rangle''' &= U^\dagger_{\mathbf{J}}(\gamma, \beta, \alpha) \,|\, j, m \,\rangle \\ &= \sum_{m'=-j}^{+j} |\, j, m' \,\rangle \,\langle\, j, m' \,|\, U^\dagger_{\mathbf{J}}(\gamma, \beta, \alpha) \,|\, j, m \,\rangle = \sum_{m'=-j}^{+j} D^{(j)\,*}_{m,m'}(\gamma, \beta, \alpha) \,|\, j, m' \,\rangle\,. \end{aligned} \tag{21.148}$$

where the D-matrix is defined by:

$$D^{(j)}_{m,m'}(\gamma,\beta,\alpha) = \langle\, j,m \,|\, U_{\mathbf{J}}(\gamma,\beta,\alpha) \,|\, j,m' \,\rangle = \langle\, j,m \,|\, e^{iJ_z\gamma/\hbar}\, e^{iJ_y\beta/\hbar}\, e^{iJ_z\alpha/\hbar} \,|\, j,m' \,\rangle \tag{21.149}$$

We warn the reader that there is a great deal of confusion, especially in the early literature, concerning Euler angles and representation of rotations in quantum mechanics. From our point of view, all we need is the matrix representation provided by Eq. (21.79) and the composition rule for unitary representation of the rotation group. Our definition of the D-matrices, Eq. (21.149), agrees with the 1996 printing of Edmonds[2][Eq. (4.1.9) on p. 55]. Earlier printings of Edmonds were in error. (See the articles by Bouten [14] and Wolf [15].)

21.3.2 Properties of D-functions

Matrix elements of the rotation operator using Euler angles to define the rotation are given by:

$$D^{(j)}_{m,m'}(\gamma,\beta,\alpha) = \langle\, jm \,|\, U_{\mathbf{J}}(\gamma,\beta,\alpha) \,|\, jm' \,\rangle = e^{i(m\gamma+m'\alpha)}\, d^{(j)}_{m,m'}(\beta)\,, \tag{21.150}$$

where $d^{j}_{m,m'}(\beta)$ is *real* and given by:[9]

$$d^{(j)}_{m,m'}(\beta) = \langle\, jm \,|\, e^{i\beta J_y/\hbar} \,|\, jm' \,\rangle\,.$$

We derive an explicit formula for the D-matrices in Theorem 72 in Section G.5 using Schwinger's methods, where we find:

$$\begin{aligned} D^{(j)}_{m,m'}(R) &= \sqrt{(j+m)!\,(j-m)!\,(j+m')\,(j-m')} \\ &\times \sum_{s=0}^{j+m}\sum_{r=0}^{j-m} \delta_{s-r,m-m'}\, \frac{\big(\,D_{+,+}(R)\,\big)^{j+m-s}\,\big(\,D_{+,-}(R)\,\big)^{s}\,\big(\,D_{-,+}(R)\,\big)^{r}\,\big(\,D_{-,-}(R)\,\big)^{j-m-r}}{s!\,(j+m-s)!\,r!\,(j-m-r)!}\,, \end{aligned} \tag{21.151}$$

where elements of the matrix $D(R)$, with rows an columns labeled by $\pm$, are given by any of the parameterizations:

$$\begin{aligned} D(R) &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \cos(\theta/2) + i\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,\sin(\theta/2) \\ &= \begin{pmatrix} e^{i(+\gamma+\alpha)/2}\,\cos(\beta/2) & e^{i(+\gamma-\alpha)/2}\,\sin(\beta/2) \\ -e^{i(-\gamma+\alpha)/2}\,\sin(\beta/2) & e^{i(-\gamma-\alpha)/2}\,\cos(\beta/2) \end{pmatrix}. \end{aligned} \tag{21.152}$$

Using Euler angles, this gives the formula:

$$\begin{aligned} d^{(j)}_{m,m'}(\beta) &= \sqrt{(j+m)!\,(j-m)!\,(j+m')\,(j-m')} \\ &\times \sum_{\sigma} \frac{(-)^{j-\sigma-m}\,\big(\,\cos(\beta/2)\,\big)^{2\sigma+m+m'}\,\big(\,\sin(\beta/2)\,\big)^{2j-2\sigma-m-m'}}{\sigma!\,(j-\sigma-m)!\,(j-\sigma-m')!\,(\sigma+m+m')!}\,. \end{aligned} \tag{21.153}$$

From this, it is easy to show that:

$$d^{(j)}_{m,m'}(\beta) = d^{(j)\,*}_{m,m'}(\beta) = d^{(j)}_{m',m}(-\beta) = (-)^{m-m'}\, d^{(j)}_{-m,-m'}(\beta) = (-)^{m-m'}\, d^{(j)}_{m',m}(\beta)\,. \tag{21.154}$$

In particular, in Section G.5, we show that:

$$d^{(j)}_{m,m'}(\pi) = (-)^{j-m}\,\delta_{m,-m'}\,, \qquad \text{and} \qquad d^{(j)}_{m,m'}(-\pi) = (-)^{j+m}\,\delta_{m,-m'}\,. \tag{21.155}$$

[9] This is the reason in quantum mechanics for choosing the second rotation to be about the y-axis rather than the x-axis.

The D-matrix for the inverse transformation is given by:

$$D^{(j)}_{m,m'}(R^{-1}) = D^{(j)*}_{m',m}(R) = (-)^{m-m'}\, D^{(j)}_{-m,-m'}(R) \tag{21.156}$$

For Euler angles, since $d^{(j)}_{m,m'}(\beta)$ is real, this means that:

$$D^{(j)*}_{m,m'}(\alpha,\beta,\gamma) = D^{(j)}_{m',m}(-\gamma,-\beta,-\alpha) = D^{(j)}_{m,m'}(-\alpha,\beta,-\gamma) = (-)^{m-m'}\, D^{(j)}_{-m,-m'}(\alpha,\beta,\gamma)\,. \tag{21.157}$$

Exercise 63. Show that the matrix $d^{(1)}(\beta)$ for $j = 1/2$, is given by:

$$d^{(1/2)}(\beta) = e^{i\beta\sigma_y/2} = \cos(\beta/2) + i\sigma_y \sin(\beta/2) = \begin{pmatrix} \cos(\beta/2) & \sin(\beta/2) \\ -\sin(\beta/2) & \cos(\beta/2) \end{pmatrix}\,,$$

so that

$$D^{(1/2)}(\gamma,\beta,\alpha) = \begin{pmatrix} e^{i(+\gamma+\alpha)/2}\cos(\beta/2) & e^{i(+\gamma-\alpha)/2}\sin(\beta/2) \\ -e^{i(-\gamma+\alpha)/2}\sin(\beta/2) & e^{i(-\gamma-\alpha)/2}\cos(\beta/2) \end{pmatrix}\,, \tag{21.158}$$

which agrees with Eq. (**??**) if we put $\gamma = 0$, $\beta = \theta$, and $\alpha = \phi$.

Exercise 64. Show that the matrix $d^{(1)}(\beta)$ for $j = 1$, is given by:

$$d^{(1)}(\beta) = e^{i\beta S_y} = \begin{pmatrix} (1+\cos\beta)/2 & \sin\beta/\sqrt{2} & (1-\cos\beta)/2 \\ -\sin\beta/\sqrt{2} & \cos\beta & \sin\beta/\sqrt{2} \\ (1-\cos\beta)/2 & -\sin\beta/\sqrt{2} & (1+\cos\beta)/2 \end{pmatrix}\,. \tag{21.159}$$

Use the results for S_y in Eq. (21.10) and expand the exponent in a power series in $i\beta S_y$ for a few terms (about four or five terms should do) in order to deduce the result directly.

Remark 37. From the results in Eq. (21.159), we note that:

$$Y_{1,m}(\theta,\phi) = \sqrt{\frac{3}{4\pi}} \begin{cases} -\sin\theta\, e^{+i\phi}/\sqrt{2}\,, & \text{for } m = +1, \\ \cos\theta\,, & \text{for } m = 0, \\ +\sin\phi\, e^{-i\phi}/\sqrt{2}\,, & \text{for } m = -1. \end{cases} \tag{21.160}$$

so

$$D^{(1)}_{0,m}(\gamma,\beta,\alpha) = \sqrt{\frac{4\pi}{3}}\, Y_{1,m}(\beta,\alpha)\,, \qquad \text{and} \qquad D^{(1)}_{m,0}(\gamma,\beta,\alpha) = (-)^m \sqrt{\frac{4\pi}{3}}\, Y_{1,m}(\beta,\gamma)\,, \tag{21.161}$$

in agreement with Eqs. (21.171) and (21.172).

21.3.3 Rotation of orbital angular momentum

When the angular momentum has a coordinate representation so that $\mathbf{J} = \mathbf{L} = \mathbf{R} \times \mathbf{P}$,

$$U^\dagger_{\mathbf{L}}(\gamma,\beta,\alpha)\, X_i\, U_{\mathbf{L}}(\gamma,\beta,\alpha) = R_{ij}(\gamma,\beta,\alpha)\, X_j = X'''_i\,, \tag{21.162}$$

or

$$X_i\, U_{\mathbf{L}}(\gamma,\beta,\alpha) = U_{\mathbf{L}}(\gamma,\beta,\alpha)\, X'''_i\,, \tag{21.163}$$

so that:

$$X_i \left\{ U_{\mathbf{L}}(\gamma,\beta,\alpha)\,|\,\mathbf{r}\,\rangle \right\} = U_{\mathbf{L}}(\gamma,\beta,\alpha)\, X'''_i\,|\,\mathbf{r}\,\rangle = x'''_i \left\{ U_{\mathbf{L}}(\gamma,\beta,\alpha)\,|\,\mathbf{r}\,\rangle \right\}, \tag{21.164}$$

which means that $U_{\mathbf{L}}(\gamma,\beta,\alpha)\,|\,\mathbf{r}\,\rangle$ is an eigenvector of X_i with eigenvalue $x'''_i = R_{ij}(\gamma,\beta,\alpha)\, x_j$. That is:

$$|\,\mathbf{r}'''\,\rangle = U_{\mathbf{L}}(\gamma,\beta,\alpha)\,|\,\mathbf{r}\,\rangle\,. \tag{21.165}$$

The spherical harmonics of Section 21.1.2 are defined by:

$$Y_{\ell,m}(\theta,\phi) = \langle\, \hat{\mathbf{r}} \,|\, \ell, m \,\rangle = \langle\, \theta, \phi \,|\, \ell, m \,\rangle \,. \tag{21.166}$$

Now let the point P be on the unit circle so that the coordinates of this point is described by the polar angles (θ, ϕ) in frame Σ and the polar angles (θ', ϕ') in the rotated frame Σ'. So on this unit circle,

$$\begin{aligned} Y_{\ell,m}(\theta,\phi) &= \langle\, \theta, \phi \,|\, \ell, m \,\rangle = \langle\, \theta''', \phi''' \,|\, U_{\mathbf{L}}(\gamma,\beta,\alpha) \,|\, \ell, m \,\rangle = \langle\, \theta''', \phi''' \,|\, \ell, m \,\rangle''' = Y'''_{\ell,m}(\theta''', \phi''') \\ &= \sum_{m'=-\ell}^{+\ell} \langle\, \theta''', \phi''' \,|\, \ell, m' \,\rangle \,\langle\, \ell, m' \,|\, U_{\mathbf{L}}(\gamma,\beta,\alpha) \,|\, \ell, m \,\rangle = \sum_{m'=-\ell}^{+\ell} Y_{\ell,m'}(\theta''', \phi''')\, D^{(\ell)}_{m',m}(\gamma,\beta,\alpha) \,, \end{aligned} \tag{21.167}$$

where

$$D^{(\ell)}_{m,m'}(\gamma,\beta,\alpha) = \langle\, \ell, m \,|\, U_{\mathbf{L}}(\gamma,\beta,\alpha) \,|\, \ell, m' \,\rangle \,. \tag{21.168}$$

As a special case, let us evaluate Eq. (21.167) at a point $P_0 = (x''', y''', z''') = (0, 0, 1)$ on the unit circle on the z'''-axis in the Σ''', or $\theta''' = 0$. However Eq. (21.30) states that:

$$Y_{\ell,m'}(0, \phi''') = \sqrt{\frac{2\ell+1}{4\pi}}\, \delta_{m',0} \,, \tag{21.169}$$

so only the $m' = 0$ term in Eq. (21.167) contributes to the sum and so evaluated at point P_0, Eq. (21.167) becomes:

$$Y_{\ell,m}(\theta,\phi) = \sqrt{\frac{2\ell+1}{4\pi}}\, D^{(\ell)}_{0,m}(\gamma,\beta,\alpha) \,. \tag{21.170}$$

The point P in the Σ frame is given by Eqs. (21.81). So for this point, the polar angles of point P in the Σ frame are: $\theta = \beta$ and $\phi = \alpha$, and Eq. (21.170) gives the result:

$$D^{(\ell)}_{0,m}(\gamma,\beta,\alpha) = \sqrt{\frac{4\pi}{2\ell+1}}\, Y_{\ell,m}(\beta,\alpha) = C_{\ell,m}(\beta,\alpha) \,. \tag{21.171}$$

By taking the complex conjugate of this expression and using properties of the spherical harmonics, we also find:

$$D^{(\ell)}_{m,0}(\gamma,\beta,\alpha) = (-)^m \sqrt{\frac{4\pi}{2\ell+1}}\, Y_{\ell,m}(\beta,\alpha) = C^*_{\ell,-m}(\beta,\alpha) \,. \tag{21.172}$$

As a special case, we find:

$$D^{(\ell)}_{0,0}(\gamma,\beta,\alpha) = P_\ell(\cos\beta) \,, \tag{21.173}$$

where $P_\ell(\cos\beta)$ is the Lagrendre polynomial of order ℓ.

Exercise 65. Prove Eq. (21.172).

21.3.4 Sequential rotations

From the general properties of the rotation group, we know that $U(R'R) = U(R')\,U(R)$. If we describe the rotations by Euler angles, we write the combined rotation as:

$$R(\gamma'', \beta'', \alpha'') = R(\gamma', \beta', \alpha')\, R(\gamma, \beta, \alpha) \,. \tag{21.174}$$

The unitary operator for this sequential transformation is then given by:

$$U_{\mathbf{J}}(\gamma'', \beta'', \alpha'') = U_{\mathbf{J}}(\gamma', \beta', \alpha')\, U_{\mathbf{J}}(\gamma, \beta, \alpha) \,. \tag{21.175}$$

So the D-functions for this sequential rotation is given by matrix elements of this expression:

$$D^{(j)}_{m,m''}(\gamma'',\beta'',\alpha'') = \sum_{m'=-j}^{+j} D^{(j)}_{m,m'}(\gamma',\beta',\alpha')\, D^{(j)}_{m',m''}(\gamma,\beta,\alpha)\,. \tag{21.176}$$

We can derive the addition theorem for spherical harmonics by considering the sequence of transformations given by:

$$R(\gamma'',\beta'',\alpha'') = R(\gamma',\beta',\alpha')\, R^{-1}(\gamma,\beta,\alpha) = R(\gamma',\beta',\alpha')\, R(-\alpha,-\beta,-\gamma)\,. \tag{21.177}$$

The D-functions for this sequential rotation for integer $j=\ell$, is given by:

$$D^{(\ell)}_{m,m''}(\gamma'',\beta'',\alpha'') = \sum_{m'=-\ell}^{+\ell} D^{(\ell)}_{m,m'}(\gamma',\beta',\alpha')\, D^{(\ell)}_{m',m''}(-\alpha,-\beta,-\gamma)\,. \tag{21.178}$$

Next, we evaluate Eq. (21.178) for $m=m''=0$. Using Eqs. (21.171), (21.172), and (21.173), we find:

$$P_\ell(\cos\beta'') = \frac{4\pi}{2\ell+1}\sum_{m'=-\ell}^{+\ell} Y_{\ell,m}(\beta',\alpha')\, Y^*_{\ell,m}(\beta,\alpha)\,. \tag{21.179}$$

Here (β,α) and (β',α') are the polar angles of two points on the unit circle in a *fixed* coordinate frame. In order to find $\cos\beta''$, we need to multiply out the rotation matrices given in Eq. (21.177). Let us first set $(\beta,\alpha)=(\theta,\phi)$ and $(\beta',\alpha')=(\theta',\phi')$, and set γ and γ' to zero. Then we find:

$$\begin{aligned} R(\gamma'',\beta'',\alpha'') &= R_y(\theta')\,R_z(\phi')\,R_z(-\phi)\,R_y(-\theta) = R_y(\theta')\,R_z(\phi'-\phi)\,R_y(-\theta) \\ &= \begin{pmatrix} \cos\theta' & 0 & -\sin\theta' \\ 0 & 1 & 0 \\ \sin\theta' & 0 & \cos\theta' \end{pmatrix} \begin{pmatrix} \cos\phi'' & \sin\phi'' & 0 \\ -\sin\phi'' & \cos\phi'' & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \sin\theta\sin\theta' + \cos\theta\cos\theta'\cos\phi'' & \cos\theta'\sin\phi'' & -\sin\theta'\cos\theta + \cos\theta'\cos\theta\cos\phi'' \\ -\cos\theta\sin\phi'' & \cos\phi'' & -\sin\theta\sin\phi'' \\ -\cos\theta'\sin\theta + \sin\theta'\cos\theta\cos\phi'' & \sin\theta'\sin\phi'' & \cos\theta'\cos\theta + \sin\theta'\sin\theta\cos\phi'' \end{pmatrix}. \end{aligned} \tag{21.180}$$

where we have set $\phi''=\phi'-\phi$. We compare this with the general form of the rotation matrix given in Eq. (21.79):

$$R(\gamma'',\beta'',\alpha'') = \begin{pmatrix} \cos\gamma''\cos\beta''\cos\alpha'' - \sin\gamma''\sin\alpha'', & \cos\gamma''\cos\beta''\sin\alpha'' + \sin\gamma''\sin\alpha'', & -\cos\gamma''\sin\beta'' \\ -\sin\gamma''\cos\beta''\cos\alpha'' - \cos\gamma''\sin\alpha'', & -\sin\gamma''\cos\beta''\sin\alpha'' + \cos\gamma''\cos\alpha'', & \sin\gamma''\sin\beta'' \\ \sin\beta''\cos\alpha'', & \sin\beta''\sin\alpha'', & \cos\beta'' \end{pmatrix}. \tag{21.181}$$

Comparing this with Eq. (21.180), we see that the $(3,3)$ component requires that:

$$\cos\beta'' = \cos\theta'\cos\theta + \sin\theta'\sin\theta\cos\phi''\,. \tag{21.182}$$

It is not easy to find the values of α'' and γ''. We leave this problem to the interested reader.

Exercise 66. Find α'' and γ'' by comparing Eqs. (21.180) and (21.181), using the result (21.182).

So Eq. (21.179) becomes:

$$P_\ell(\cos\gamma) = \frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{+\ell} Y_{\ell,m}(\theta',\phi')\, Y^*_{\ell,m}(\theta,\phi)\,, \tag{21.183}$$

where $\cos\gamma = \cos\theta'\cos\theta + \sin\theta'\sin\theta\cos(\phi'-\phi)$. Eq. (21.183) is called the addition theorem of spherical harmonics.

21.4 Addition of angular momentum

If a number of angular momentum vectors commute, the eigenvectors of the combined system can be written as a direct product consisting of the vectors of each system:

$$|\, j_1, m_1, j_2, m_2, \ldots, j_N, m_N \,\rangle = |\, j_1, m_1 \,\rangle \otimes |\, j_2, m_2 \,\rangle \otimes \cdots \otimes |\, j_N, m_N \,\rangle \,. \tag{21.184}$$

This vector is an eigenvector of J_i^2 and $J_{i,z}$ for $i = 1, 2, \ldots, N$. It is also an eigenvector of the total z-component of angular momentum: $J_z \,|\, j_1, m_1, j_2, m_2, \ldots, j_N, m_N \,\rangle = M \,|\, j_1, m_1, j_2, m_2, \ldots, j_N, m_N \,\rangle$, where $M = m_1 + m_2 + \cdots + m_N$. It is *not*, however, an eigenvector of the total angular momentum J^2, defined by

$$J^2 = \mathbf{J} \cdot \mathbf{J} \,, \qquad \mathbf{J} = \sum_{i=1}^{N} \mathbf{J}_i \,. \tag{21.185}$$

We can find eigenvectors of the total angular momentum of any number of commuting angular momentum vectors by coupling them in a number of ways. This coupling is important in applications since very often the *total* angular momentum of a system is conserved. We show how to do this coupling in this section. We start with the coupling of the eigenvectors of two angular momentum vectors.

21.4.1 Coupling of two angular momenta

Let $\mathbf{J}_1$ and $\mathbf{J}_2$ be two commuting angular moment vectors: $[\, J_{1\,i}, J_{2\,j} \,] = 0$, with $[\, J_{1\,i}, J_{1\,j} \,] = i\epsilon_{ijk} J_{1\,k}$ and $[\, J_{2\,i}, J_{12,j} \,] = i\epsilon_{ijk} J_{2\,k}$. One set of four commuting operators for the combined system is the direct product set, given by: $(J_1^2, J_{1\,z}, J_2^2, J_{2,z})$, and with eigenvectors:

$$|\, j_1, m_1, j_2, m_2 \,\rangle \,. \tag{21.186}$$

However, we can find another set of four commuting operators by defining the total angular momentum operator:

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 \,, \tag{21.187}$$

which obeys the usual angular momentum commutation rules: $[\, J_i, J_j \,] = i\epsilon_{ijk} J_k$, with $[\, J^2, J_1^2 \,] = [\, J^2, J_2^2 \,] = 0$. So another set of four commuting operators for the combined system is: (J_1^2, J_2^2, J^2, J_z), with eigenvectors:

$$|\, (j_1, j_2), j, m \,\rangle \,. \tag{21.188}$$

Either set of eigenvectors are equivalent descriptions of the combined angular momentum system, and so there us a unitary operator relating them. Matrix elements of this operator are called **Clebsch-Gordan** coefficients, or vector coupling coefficients, which we write as:

$$|\, (j_1, j_2)\, j, m \,\rangle = \sum_{m_1, m_2} |\, j_1, m_1, j_2, m_2 \,\rangle \,\langle\, j_1, m_1, j_2, m_2 \,|\, (j_1, j_2)\, j, m \,\rangle \,, \tag{21.189}$$

or in the reverse direction:

$$|\, j_1, m_1, j_2, m_2 \,\rangle = \sum_{j,m} |\, (j_1, j_2), j, m \,\rangle \langle\, (j_1, j_2)\, j, m \,|\, j_1, m_1, j_2, m_2 \,\rangle \,. \tag{21.190}$$

Since the basis states are orthonormal and complete, Clebsch-Gordan coefficients satisfy:

$$\begin{aligned} \sum_{m_1, m_2} \langle\, (j_1, j_2)\, j, m \,|\, j_1, m_1, j_2, m_2 \,\rangle \,\langle\, j_1, m_1, j_2, m_2 \,|\, (j_1, j_2)\, j', m' \,\rangle &= \delta_{j,j'}\, \delta_{m,m'} \,, \\ \sum_{j,m} \langle\, j_1, m_1, j_2, m_2 \,|\, (j_1, j_2)\, j, m \,\rangle \langle\, (j_1, j_2)\, j, m \,|\, j_1, m_1', j_2, m_2' \,\rangle &= \delta_{m_1, m_1'}\, \delta_{m_2, m_2'} \,. \end{aligned} \tag{21.191}$$

In addition, a phase convension is adopted so that the phase of the Clebsch-Gordan coefficient $\langle\, j_1, j_1, j_2, j - j_1 \,|\, (j_1, j_2)\, j, m \,\rangle$ is taken to be zero, i.e. the argument is $+1$. With this convention, all Clebsch-Gordan coefficients are real.

Operating on (21.189) by $J_z = J_{1\,z} + J_{2\,z}$, gives

$$m \,|\, (j_1, j_2)\, j, m \,\rangle = \sum_{m_1, m_2} (\, m_1 + m_2 \,) \,|\, j_1, m_1, j_2, m_2 \,\rangle \,\langle\, j_1, m_1, j_2, m_2 \,|\, (j_1, j_2)\, j, m \,\rangle \,, \tag{21.192}$$

or

$$(\, m - m_1 - m_2 \,) \,\langle\, j_1, m_1, j_2, m_2 \,|\, (j_1, j_2)\, j, m \,\rangle = 0 \,, \tag{21.193}$$

so that Clebsch-Gordan coefficients vanish unless $m = m_1 + m_2$. Operating on (21.189) by $J_\pm = J_{1\,\pm} + J_{2\,\pm}$ gives two recursion relations:

$$\begin{aligned} &A(j, \mp m) \,\langle\, j_1, m_1, j_2, m_2 \,|\, (j_1, j_2)\, j, m \pm 1 \,\rangle = \\ &\quad A(j_1, \pm m_1) \,\langle\, j_1, m_1 \mp 1, j_2, m_2 \,|\, (j_1, j_2)\, j, m \,\rangle + A(j_2, \pm m_2) \,\langle\, j_1, m_1, j_2, m_2 \mp 1 \,|\, (j_1, j_2)\, j, m \,\rangle \,, \end{aligned} \tag{21.194}$$

where $A(j, m) = \sqrt{(j + m)(j - m + 1)} = A(j, 1 \mp m)$. The range of j is determined by noticing that $\langle\, j_1, m_1, j - j_1, m_2 \,|\, (j_1, j_2)\, j, m \,\rangle$ vanished unless $-j_2 \le j - j_1 \le j_2$ or $j_1 - j_2 \le j \le j_1 + j_2$. Similarly $\langle\, j_1, j - j_2, j_2, j_2 \,|\, (j_1, j_2)\, j, m \,\rangle$ vanished unless $-j_1 \le j - j_2 \le j_1$ or $j_2 - j_1 \le j \le j_1 + j_2$, from which we conclude that

$$|\, j_1 - j_2 \,| \le j \le j_1 + j_2 \,, \tag{21.195}$$

which is called the **triangle inequality**. One can find a closed form for the Clebsch-Gordan coefficients by solving the recurrence formula, Eq. (21.194). The result [5][p. 78], which is straightforward but tedious is:

$$\begin{aligned} &\langle\, j_1, m_1, j_2, m_2 \,|\, (j_1, j_2)\, j, m \,\rangle \\ &\quad = \delta_{m, m_1 + m_2} \left[\frac{(2j + 1)\,(j_1 + j_2 - j)!\,(j_1 - m_1)!\,(j_2 - m_2)!\,(j - m)!\,(j + m)!}{(j_1 + j_2 + j + 1)!\,(j + j_1 - j_2)!\,(j + j_2 - j_1)!\,(j_1 + m_1)!\,(j_2 + m_2)!} \right]^{1/2} \\ &\qquad \times \sum_t (-)^{j_1 - m_1 + t} \left[\frac{(j_1 + m_1 + t)!\,(j + j_2 - m_1 - t)!}{t!\,(j - m - t)!\,(j_1 - m_1 - t)!\,(j_2 - j + m_1 + t)!} \right] . \end{aligned} \tag{21.196}$$

This form for the Clebsch-Gordan coefficient is called "Racah's first form." A number of other forms of the equation can be obtained by substitution. For numerical calculatins for small j, it is best to start with the vector for $m = -j$ and then apply J_+ to obtain vectors for the other m-values, or start with the vector for $m = +j$ and then apply J_- to obtain vectors for the rest of the m-values. Orthonormalization requirements between states with different value of j with the same value of m can be used to further fix the vectors. We illustrate this method in the next example.

Example 33. For $j_1 = j_2 = 1/2$, the total angular momentum can have the values $j = 0, 1$. For this example, let us simplify our notation and put $|\, 1/2, m, 1/2, m' \,\rangle \mapsto |\, m, m' \,\rangle$ and $|\, (1/2, 1/2)\, j, m \,\rangle \mapsto |\, j, m \,\rangle$. Then for $j = 1$ and $m = 1$, we start with the unique state:

$$|\, 1, 1 \,\rangle = |\, 1/2, 1/2 \,\rangle \,. \tag{21.197}$$

Our convention is that the argument of this Clebsch-Gordan coefficient is $+1$. Apply J_- to this state:

$$J_- \,|\, 1, 1 \,\rangle = J_{1\,-} \,|\, 1/2, 1/2 \,\rangle + J_{2\,-} \,|\, 1/2, 1/2 \,\rangle \,, \tag{21.198}$$

from which we find:

$$|\, 1, 0 \,\rangle = \frac{1}{\sqrt{2}} \left(|\, -1/2, 1/2 \,\rangle + |\, 1/2, -1/2 \,\rangle \right) . \tag{21.199}$$

Applying J_- again to this state gives:

$$|\,1,-1\,\rangle = |\,-1/2,-1/2\,\rangle\,. \tag{21.200}$$

For the $j=0$ case, we have:

$$|\,0,0\,\rangle = \alpha\,|\,1/2,-1/2\,\rangle + \beta\,|\,-1/2,1/2\,\rangle\,. \tag{21.201}$$

Applying J_- to this state gives zero on the left-hand-side, so we find that $\beta=-\alpha$. Since our convention is that the argument of α is $+1$, we find:

$$|\,0,0\,\rangle = \frac{1}{\sqrt{2}}\left(|\,1/2,-1/2\,\rangle - |\,-1/2,1/2\,\rangle\right). \tag{21.202}$$

As a check, we note that (21.202) is orthogonal to (21.199). We summarize these familiar results as follows:

$$|\,j,m\,\rangle = \begin{cases} \left(|\,1/2,-1/2\,\rangle - |\,-1/2,1/2\,\rangle\right)/\sqrt{2}\,, & \text{for } j=m=0,\\ |\,1/2,1/2\,\rangle\,, & \text{for } j=1,\ m=+1,\\ \left(|\,1/2,-1/2\,\rangle + |\,-1/2,1/2\,\rangle\right)/\sqrt{2}\,, & \text{for } j=1,\ m=0,\\ |\,-1/2,-1/2\,\rangle\,, & \text{for } j=1,\ m=-1. \end{cases} \tag{21.203}$$

Exercise 67. Work out the Clebsch-Gordan coefficients for the case when $j_1=1/2$ and $j_2=1$.

Tables of Clebsch-Gordan coefficients can be found on the internet. We reproduce one of them from the Particle Data group in Table 21.1.[10] More extensive tables can be found in the book by Rotenberg, et.al. [16], and computer programs for numerically calculating Clebsch-Gordan coefficients, 3j-, 6j-, and 9j-symbols are also available. Important symmetry relations for Clebsch-Gordan coefficients are the following:

1. Interchange of the order of (j_1,j_2) coupling:

$$\langle\, j_2,m_2,j_1,m_1\,|\,(j_2,j_1)\,j_3,m_3\,\rangle = (-)^{j_1+j_2-j_3}\,\langle\, j_1,m_1,j_2,m_2\,|\,(j_1,j_2)\,j_3,m_3\,\rangle\,. \tag{21.204}$$

2. Cyclic permutation of the coupling $[(j_1,j_2)\,j_3]$:

$$\langle\, j_2,m_2,j_3,m_3\,|\,(j_2,j_3)\,j_1,m_1\,\rangle = (-)^{j_2-m_2}\sqrt{\frac{2j_1+1}{2j_3+1}}\,\langle\, j_1,m_1,j_2,-m_2\,|\,(j_1,j_2)\,j_3,m_3\,\rangle\,, \tag{21.205}$$

$$\langle\, j_3,m_3,j_1,m_1\,|\,(j_3,j_1)\,j_2,m_2\,\rangle = (-)^{j_1+m_1}\sqrt{\frac{2j_2+1}{2j_3+1}}\,\langle\, j_1,-m_1,j_2,m_2\,|\,(j_1,j_2)\,j_3,m_3\,\rangle\,. \tag{21.206}$$

3. Reversal of all m values:

$$\langle\, j_1,-m_1,j_2,-m_2\,|\,(j_1,j_2)\,j_3,-m_3\,\rangle = (-)^{j_1+j_2-j_3}\,\langle\, j_1,m_1,j_2,m_2\,|\,(j_1,j_2)\,j_3,m_3\,\rangle\,. \tag{21.207}$$

Some special values of the Clebsch-Gordan coefficients are useful to know:

$$\langle\, j,m,0,0\,|\,(j,0)\,j,m\,\rangle = 1\,, \qquad \langle\, j,m,j,m'\,|\,(j,j)\,0,0\,\rangle = \delta_{m,-m'}\,\frac{(-)^{j-m}}{\sqrt{2j+1}}\,. \tag{21.208}$$

The symmetry relations are most easily obtained from the simpler symmetry relations for $3j$-symbols, which are defined below, and proved in Section G.6 using Schwinger's methods.

[10] The sign convensions for the d-functions in this table are those of Rose[6], who uses an active rotation. To convert them to our conventions put $\beta\to-\beta$.

34. CLEBSCH-GORDAN COEFFICIENTS, SPHERICAL HARMONICS, AND *d* FUNCTIONS

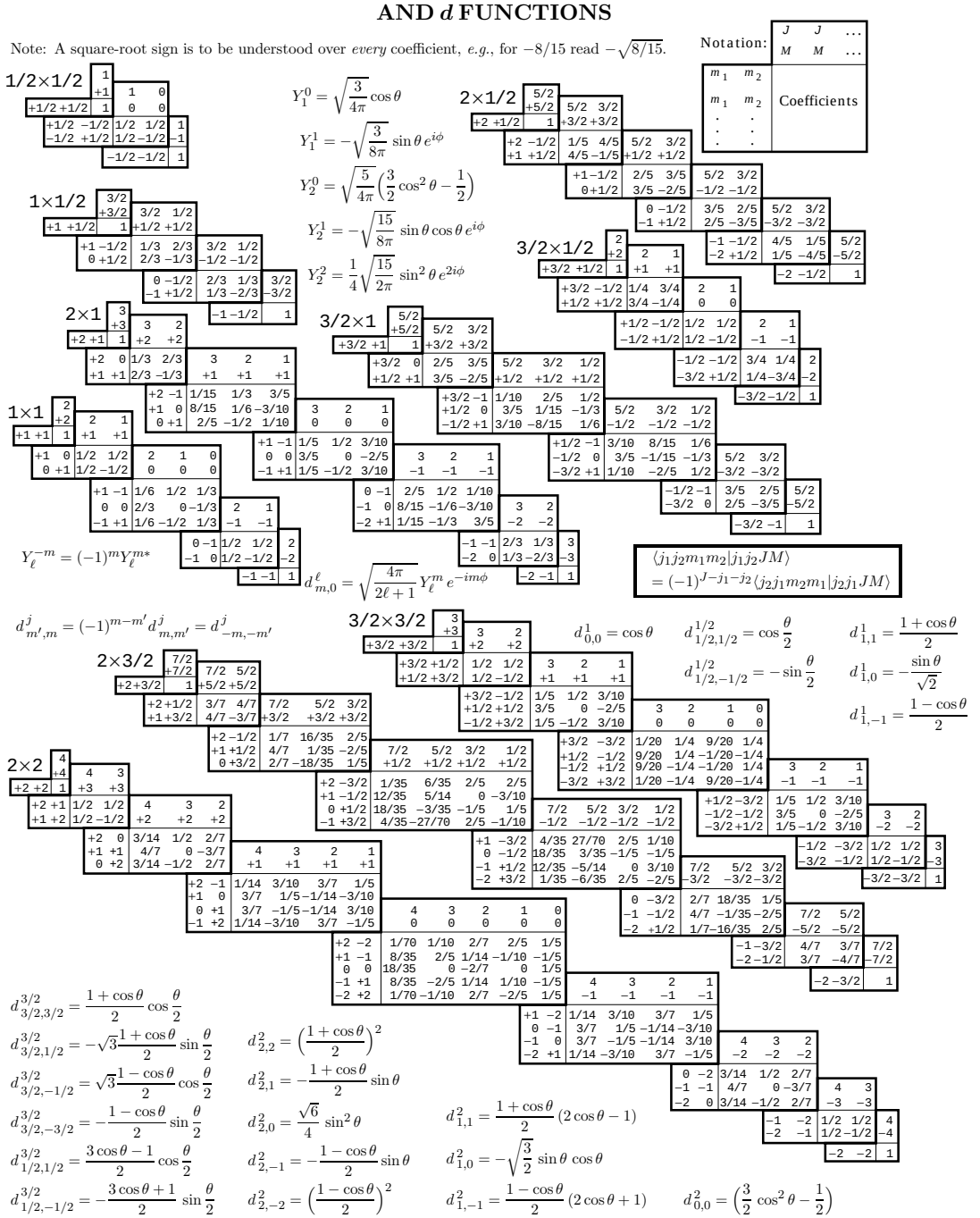

Figure 34.1: The sign convention is that of Wigner (*Group Theory*, Academic Press, New York, 1959), also used by Condon and Shortley (*The Theory of Atomic Spectra*, Cambridge Univ. Press, New York, 1953), Rose (*Elementary Theory of Angular Momentum*, Wiley, New York, 1957), and Cohen (*Tables of the Clebsch-Gordan Coefficients*, North American Rockwell Science Center, Thousand Oaks, Calif., 1974). The coefficients here have been calculated using computer programs written independently by Cohen and at LBNL.

Table 21.1: Table of Clebsch-Gordan coefficients, spherical harmonics, and *d*-functions.

3j symbols

Clebsch-Gordan coefficients do not possess simple symmetry relations upon exchange of the angular momentum quantum numbers. 3-j symbols which are related to Clebsch-Gordan coefficients, have better symmetry properties. They are defined by (Edmonds [2]):

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \frac{(-)^{j_1-j_2-m_3}}{\sqrt{2j_3+1}} \langle\, j_1, m_1, j_2, m_2 \,|\, (j_1, j_2)\, j_3, -m_3 \,\rangle \,. \tag{21.209}$$

In terms of $3j$-symbols, the orthogonality relations (21.191) become:

$$\begin{aligned} (2j_3+1) \sum_{m_1,m_2} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3' \\ m_1 & m_2 & m_3' \end{pmatrix} &= \delta_{j_3,j_3'}\, \delta_{m_3,m_3'} \,, \\ \sum_{j_3,m_3} (2j_3+1) \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1' & m_2' & m_3 \end{pmatrix} &= \delta_{m_1,m_1'}\, \delta_{m_2,m_2'} \,. \end{aligned} \tag{21.210}$$

Symmetry properties of the $3j$-symbols are particularly simple. They are:

1. The $3j$-symbols are invariant under even (cyclic) permutation of the columns:

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} j_2 & j_3 & j_1 \\ m_2 & m_3 & m_1 \end{pmatrix} = \begin{pmatrix} j_3 & j_1 & j_2 \\ m_3 & m_1 & m_2 \end{pmatrix} \,. \tag{21.211}$$

   and are multiplied by a phase for odd permutations:

$$\begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} = \begin{pmatrix} j_3 & j_2 & j_1 \\ m_3 & m_2 & m_1 \end{pmatrix} = \begin{pmatrix} j_1 & j_3 & j_2 \\ m_1 & m_3 & m_2 \end{pmatrix} = (-)^{j_1+j_2+j_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \,. \tag{21.212}$$

2. For reversal of all m values:

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} = (-)^{j_1+j_2+j_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \,. \tag{21.213}$$

The $3j$-symbols vanish unless $m_1 + m_2 + m_3 = 0$. For $j_3 = 0$, the $3j$-symbol is:

$$\begin{pmatrix} j & j & 0 \\ m & m' & 0 \end{pmatrix} = \delta_{m,-m'}\, \frac{(-)^{j-m}}{\sqrt{2j+1}} \,. \tag{21.214}$$

A few useful $3j$-symbols are given in Table 21.2. More can be found in Edmonds [2][Table 2, p. 125] and Brink and Satchler [17][Table 3, p. 36].

21.4.2 Coupling of three and four angular momenta

We write the direct product eigenvector for three angular momenta as:

$$|\, j_1, m_1, j_2, m_2, j_3, m_3 \,\rangle = |\, j_1, m_1 \,\rangle \otimes |\, j_2, m_2 \,\rangle \otimes |\, j_3, m_3 \,\rangle \,. \tag{21.215}$$

This state is an eigenvector of $J_1^2, J_{1\,z}$, $J_2^2, J_{2\,z}$, and $J_3^2, J_{3\,z}$. If we want to construct eigenvectors of total angular momentum J^2 and J_z, where

$$J^2 = \mathbf{J} \cdot \mathbf{J} \,, \qquad \mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 + \mathbf{J}_3 \,, \tag{21.216}$$

there are three ways to do this: (1) couple $\mathbf{J}_1$ to $\mathbf{J}_2$ to get an intermediate vector $\mathbf{J}_{12}$ and then couple this intermediate vector to $\mathbf{J}_3$ to get an eigenvector of $\mathbf{J}$, (2) couple $\mathbf{J}_2$ to $\mathbf{J}_3$ to get $\mathbf{J}_{23}$ and then couple $\mathbf{J}_1$ to

$$\begin{aligned}
\begin{pmatrix} j & j+1/2 & 1/2 \\ m & -m-1/2 & 1/2 \end{pmatrix} &= (-)^{j-m-1}\sqrt{\frac{j+m+1}{(2j+1)(2j+2)}} \\
\begin{pmatrix} j & j & 1 \\ m & -m-1 & 1 \end{pmatrix} &= (-)^{j-m}\sqrt{\frac{2(j-m)(j+m+1)}{2j(2j+1)(2j+2)}} \\
\begin{pmatrix} j & j & 1 \\ m & -m & 0 \end{pmatrix} &= (-)^{j-m}\frac{m}{\sqrt{j(j+1)(2j+1)}} \\
\begin{pmatrix} j & j+1 & 1 \\ m & -m-1 & 1 \end{pmatrix} &= (-)^{j-m}\sqrt{\frac{(j+m+1)(j+m+2)}{(2j+1)(2j+2)(2j+3)}} \\
\begin{pmatrix} j & j+1 & 1 \\ m & -m & 0 \end{pmatrix} &= (-)^{j-m-1}\sqrt{\frac{2(j-m+1)(j+m+1)}{(2j+1)(2j+2)(2j+3)}}
\end{aligned}$$

Table 21.2: Algebric formulas for some $3j$-symbols.

$\mathbf{J}_{23}$ to get $\mathbf{J}$, or (3) couple $\mathbf{J}_1$ to $\mathbf{J}_3$ to get $\mathbf{J}_{13}$ and then couple $\mathbf{J}_2$ to $\mathbf{J}_{13}$ to get $\mathbf{J}$. Keeping in mind that the *order* of the coupling of two vectors is just a phase and not a different coupling scheme, it turns out that this last coupling is just a combined transformation of the first two (see Eq. (21.223) below). So the first two coupling schemes can be written as:

$$\begin{aligned}
|\,(j_1,j_2)\,j_{12},j_3,j,m\,\rangle &= \sum_{\substack{m_1,m_2,m_3\\ m_{12}}} \langle\, j_1,m_1,j_2,m_2\,|\,(j_1,j_2)\,j_{12},m_{12}\,\rangle\,\langle\, j_{12},m_{12},j_3,m_3\,|\,(j_{12},j_3)\,j,m\,\rangle \\
&\qquad\qquad \times |\,j_1,m_1,j_2,m_2,j_3,m_3\,\rangle\,, \\
|\,j_1\,(j_2,j_3)\,j_{23},j,m\,\rangle &= \sum_{\substack{m_1,m_2,m_3\\ m_{23}}} \langle\, j_2,m_2,j_3,m_3\,|\,(j_2,j_3)\,j_{23},m_{23}\,\rangle\,\langle\, j_1,m_1,j_{23},m_{23}\,|\,(j_1,j_{23})\,j,m\,\rangle \\
&\qquad\qquad \times |\,j_1,m_1,j_2,m_2,j_3,m_3\,\rangle\,.
\end{aligned} \tag{21.217}$$

The overlap between these two coupling vectors is independent of m and is proportional to the $6j$-symbol:

$$\begin{aligned}
\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix} &= \frac{(-)^{j_1+j_2+j_3+j}}{\sqrt{(2j_{12}+1)\,(2j_{23}+1)}}\,\langle\,(j_1,j_2)\,j_{12},j_3,j,m\,|\,j_1\,(j_2,j_3)\,j_{23},j,m\,\rangle \\
&= \frac{(-)^{j_1+j_2+j_3+j}}{\sqrt{(2j_{12}+1)\,(2j_{23}+1)}}\sum_{\substack{m_1,m_2,m_3\\ m_{12},m_{23}}} \langle\, j_1,m_1,j_2,m_2\,|\,(j_1,j_2)\,j_{12},m_{12}\,\rangle
\end{aligned} \tag{21.218}$$

$$\times\,\langle\, j_{12},m_{12},j_3,m_3\,|\,(j_{12},j_3)\,j,m\,\rangle\,\langle\, j_2,m_2,j_3,m_3\,|\,(j_2,j_3)\,j_{23},m_{23}\,\rangle\,\langle\, j_1,m_1,j_{23},m_{23}\,|\,(j_1,j_{23})\,j,m\,\rangle$$

Here $m = m_1 + m_2 + m_3$. The $6j$-symbols vanish unless (j_1,j_2,j_{12}), (j_2,j_3,j_{23}), (j_{12},j_3,j), and (j_1,j_{23},j) all satisfy triangle inequalities. In terms of $3j$-symbols, the $6j$-symbol is given by the symmetric expression:

$$\begin{aligned}
\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{Bmatrix} &= \sum_{\text{all } m} (-)^p \\
&\times \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}\begin{pmatrix} j_1 & j_5 & j_6 \\ -m_1 & m_5 & -m_6 \end{pmatrix}\begin{pmatrix} j_4 & j_2 & j_6 \\ -m_4 & -m_2 & m_6 \end{pmatrix}\begin{pmatrix} j_4 & j_5 & j_3 \\ m_4 & -m_5 & -m_3 \end{pmatrix},
\end{aligned} \tag{21.219}$$

where $p = j_1+j_2+j_3+j_4+j_5+j_6+m_1+m_2+m_3+m_4+m_5+m_6$. Here, the sums over all m's are restructed because the $3j$-symbols vanish unless their m-values add to zero. A number of useful relations between $3j$ and $6j$-symbols follow from Eq. (21.219), and are tablulated by Brink and Satchler [17][Appendix II, p. 141]. One of these which we will use later is:

$$\sqrt{(2\ell+1)(2\ell'+1)}\begin{Bmatrix} \ell & \ell' & k \\ j' & j & 1/2 \end{Bmatrix}\begin{pmatrix} \ell & \ell' & k \\ 0 & 0 & 0 \end{pmatrix} = (-)^{j+\ell+j'+\ell'+1}\begin{pmatrix} \ell' & \ell & k \\ -1/2 & 1/2 & 0 \end{pmatrix}\delta(\ell,\ell',k)\,, \tag{21.220}$$

$$
\begin{aligned}
\begin{Bmatrix} j_1 & j_2 & j_3 \\ 0 & j_3 & j_2 \end{Bmatrix} &= \frac{(-)^{j_1+j_2+j_3}}{\sqrt{(2j_2+1)(2j_3+1)}}\,, \\
\begin{Bmatrix} j_1 & j_2 & j_3 \\ 1/2 & j_3-1/2 & j_2+1/2 \end{Bmatrix} &= (-)^{j_1+j_2+j_3}\sqrt{\frac{(j_1+j_3-j_2)(j_1+j_2-j_3+1)}{(2j_2+1)(2j_2+2)2j_3(2j_3+1)}}\,, \\
\begin{Bmatrix} j_1 & j_2 & j_3 \\ 1/2 & j_3-1/2 & j_2-1/2 \end{Bmatrix} &= (-)^{j_1+j_2+j_3}\sqrt{\frac{(j_2+j_3-j_1)(j_1+j_2+j_3+1)}{2j_2(2j_2+1)2j_3(2j_3+1)}}\,, \\
\begin{Bmatrix} j_1 & j_2 & j_3 \\ 1 & j_3 & j_2 \end{Bmatrix} &= 2\,(-)^{j_1+j_2+j_3}\,\frac{j_1(j_1+1)-j_2(j_2+1)-j_3(j_3+1)}{\sqrt{2j_2(2j_2+1)(2j_2+2)\,2j_3(2j_3+1)(2j_3+2)}}\,,
\end{aligned}
$$

Table 21.3: Algebric formulas for some $6j$-symbols.

where $\delta(\ell,\ell',k)=1$ if $\ell+\ell'+k$ is even and (ℓ,ℓ',k) satisfy the triangle inequality, otherwise it is zero. The $6j$-symbol is designed so as to maximize the symmetries of the coupling coefficient, as in the $3j$-symbol. For example, the $6j$-symbol is invariant under any permutation of columns:

$$
\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{Bmatrix} = \begin{Bmatrix} j_2 & j_3 & j_1 \\ j_5 & j_6 & j_4 \end{Bmatrix} = \begin{Bmatrix} j_3 & j_1 & j_2 \\ j_6 & j_4 & j_5 \end{Bmatrix} = \begin{Bmatrix} j_2 & j_1 & j_3 \\ j_5 & j_4 & j_6 \end{Bmatrix} = \begin{Bmatrix} j_1 & j_3 & j_2 \\ j_4 & j_6 & j_5 \end{Bmatrix} = \begin{Bmatrix} j_3 & j_2 & j_1 \\ j_6 & j_5 & j_4 \end{Bmatrix}\,.
$$

It is also invariant under exchange of the upper and lower elements of any two columns:

$$
\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{Bmatrix} = \begin{Bmatrix} j_4 & j_5 & j_3 \\ j_1 & j_2 & j_6 \end{Bmatrix} = \begin{Bmatrix} j_4 & j_2 & j_6 \\ j_1 & j_5 & j_3 \end{Bmatrix} = \begin{Bmatrix} j_1 & j_5 & j_6 \\ j_4 & j_2 & j_3 \end{Bmatrix}\,.
$$

Some particular $6j$-symbols are given in Table 21.3 Additional tables of $6j$-symbols for values of $j=1$ and 2 can be found in Edmonds [2][Table 5, p. 130]. Several relations between $6j$-symbols are obtained by consideration of the recoupling matrix elements. For example, since:

$$
\sum_{j_{12}} \langle\, j_1\,(j_2,j_3)\,j_{23}, j\,|\,(j_1,j_2)\,j_{12}, j_3, j\,\rangle\,\langle\,(j_1,j_2)\,j_{12}, j_3, j\,|\,j_1\,(j_2,j_3)\,j'_{23}, j\,\rangle = \delta_{j_{23},j'_{23}}\,, \tag{21.221}
$$

we have:

$$
\sum_{j_{12}} (2j_{12}+1)(2j_{23}+1)\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix}\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j'_{23} \end{Bmatrix} = \delta_{j_{23},j'_{23}}\,. \tag{21.222}
$$

A similar consideration of

$$
\begin{aligned}
\sum_{j_{23}} &\langle\,(j_1,j_2)\,j_{12}, j_3, j\,|\,j_1\,(j_2,j_3)\,j_{23}, j\,\rangle\,\langle\,j_1\,(j_2,j_3)\,j_{23}, j\,|\,j_2\,(j_3,j_1)\,j_{31}, j\,\rangle \\
&= \langle\,(j_1,j_2)\,j_{12}, j_3, j\,|\,j_2\,(j_3,j_1)\,j_{31}, j\,\rangle\,,
\end{aligned} \tag{21.223}
$$

gives:

$$
\sum_{j_{23}} (-)^{j_{23}+j_{31}+j_{12}}\,(2j_{23}+1)\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix}\begin{Bmatrix} j_2 & j_3 & j_{23} \\ j_1 & j & j_{31} \end{Bmatrix} = \begin{Bmatrix} j_3 & j_1 & j_{31} \\ j_2 & j & j_{12} \end{Bmatrix}\,. \tag{21.224}
$$

Other important formula involving $6j$-symbols can be found in standard references.

The coupling of four angular momenta is done in a similar way. Let us take the special case of two particles with orbital angular momentum ℓ_1 and ℓ_2 and spin s_1 and s_2. Two important ways of coupling these four angular momentum are the j-j coupling scheme:

$$
\begin{aligned}
&|\,(\ell_1,s_1)\,j_1, (\ell_2,s_2)\,j_2, j, m\,\rangle = \\
&\sum_{\substack{m_{\ell_1},m_{s_1},m_{\ell_2},m_{s_2} \\ m_{j_1},m_{j_2}}} \langle\,\ell_1,m_{\ell_1},s_1,m_{s_1}\,|\,(\ell_1,s_1)\,j_1,m_1\,\rangle\langle\,\ell_2,m_{\ell_2},s_2,m_{s_2}\,|\,(\ell_2,s_2)\,j_2,m_2\,\rangle\,\langle\,j_1,m_1,j_2,m_2\,|\,(j_1,j_2)\,j,m\,\rangle,
\end{aligned} \tag{21.225}
$$

and the ℓ-s coupling scheme:

$$|\,(\ell_1,\ell_2)\,\ell,(s_1,s_2)\,s,j,m\,\rangle = \sum_{\substack{m_{\ell_1},m_{\ell_2},m_{s_1},m_{s_2}\\ m_\ell,m_s}} \langle\,\ell_1,m_{\ell_1},\ell_2,m_{\ell_2}\,|\,(\ell_1,\ell_2)\,\ell,m_\ell\,\rangle\,\langle\,s_1,m_{s_1},s_2,m_{s_2}\,|\,(s_1,s_2)\,s,m_s\,\rangle\,\langle\,\ell,m_\ell,s,m_s\,|\,(\ell,s)\,j,m\,\rangle\,, \tag{21.226}$$

The overlap between these two coupling schemes define the $9j$-symbol:

$$\begin{Bmatrix}\ell_1 & s_1 & j_1\\ \ell_2 & s_2 & j_2\\ \ell & s & j\end{Bmatrix} = \frac{\langle\,(\ell_1,s_1)\,j_1,(\ell_2,s_2)\,j_2,j,m\,|\,(\ell_1,\ell_2)\,\ell,(s_1,s_2)\,s,j,m\,\rangle}{\sqrt{(2j_1+1)\,(2j_2+1)\,(2\ell+1)\,(2s+1)}} \tag{21.227}$$

and is independent of the value of m. The rows and columns of the $9j$-symbol must satisfy the triangle inequality. From Eqs. (21.225) and (21.226), the $9j$-symbol can be written in terms of sums over $6j$-symbols or $3j$-symbols:

$$\begin{aligned}\begin{Bmatrix}j_{11} & j_{12} & j_{13}\\ j_{21} & j_{22} & j_{23}\\ j_{31} & j_{32} & j_{33}\end{Bmatrix} &= \sum_j (-)^{2j}\,(2j+1)\begin{Bmatrix}j_{11} & j_{21} & j_{31}\\ j_{32} & j_{33} & j\end{Bmatrix}\begin{Bmatrix}j_{12} & j_{22} & j_{32}\\ j_{21} & j & j_{23}\end{Bmatrix}\begin{Bmatrix}j_{13} & j_{23} & j_{33}\\ j & j_{11} & j_{12}\end{Bmatrix}\\ &= \sum_{\text{all } m}\begin{pmatrix}j_{11} & j_{12} & j_{13}\\ m_{11} & m_{12} & m_{13}\end{pmatrix}\begin{pmatrix}j_{21} & j_{22} & j_{23}\\ m_{21} & m_{22} & m_{23}\end{pmatrix}\begin{pmatrix}j_{31} & j_{32} & j_{33}\\ m_{31} & m_{32} & m_{33}\end{pmatrix}\\ &\quad\times\begin{pmatrix}j_{11} & j_{21} & j_{31}\\ m_{11} & m_{21} & m_{31}\end{pmatrix}\begin{pmatrix}j_{12} & j_{22} & j_{32}\\ m_{12} & m_{22} & m_{32}\end{pmatrix}\begin{pmatrix}j_{13} & j_{23} & j_{33}\\ m_{13} & m_{23} & m_{33}\end{pmatrix}.\end{aligned} \tag{21.228}$$

From Eq. (21.228), we see that an even permutation of rows or columns or a transposition of rows and columns leave the $9j$-symbol invariant, whereas an odd permutation of rows or columns produces a sign change given by:

$$(-)^{j_{11}+j_{12}+j_{13}+j_{21}+j_{22}+j_{23}+j_{31}+j_{32}+j_{33}}\,.$$

Orthogonal relations of $9j$-symbols are obtained in the same way as with the $3j$-symbols. We find:

$$\sum_{j_{12},j_{34}}(2j_{12}+1)(2j_{34}+1)(2j_{13}+1)(2j_{24}+1)\begin{Bmatrix}j_1 & j_2 & j_{12}\\ j_3 & j_4 & j_{34}\\ j_{13} & j_{24} & j\end{Bmatrix}\begin{Bmatrix}j_1 & j_2 & j_{12}\\ j_3 & j_4 & j_{34}\\ j'_{13} & j'_{24} & j\end{Bmatrix} = \delta_{j_{13},j'_{13}}\,\delta_{j_{24},j'_{24}}\,, \tag{21.229}$$

and

$$\begin{aligned}\sum_{j_{13},j_{24}}(-)^{2j_3+j_{24}+j_{23}-j_{34}}(2j_{13}+1)(2j_{24}+1)&\begin{Bmatrix}j_1 & j_2 & j_{12}\\ j_3 & j_4 & j_{34}\\ j_{13} & j_{24} & j\end{Bmatrix}\begin{Bmatrix}j_1 & j_3 & j_{13}\\ j_4 & j_2 & j_{24}\\ j_{14} & j_{23} & j\end{Bmatrix}\\ &= \begin{Bmatrix}j_1 & j_2 & j_{12}\\ j_4 & j_3 & j_{34}\\ j_{14} & j_{23} & j\end{Bmatrix}.\end{aligned} \tag{21.230}$$

Relations between $6j$- and $9j$-symbols are obtained from orthogonality relations and recoupling vectors. One which we will have occasion to use is:

$$\sum_{j_{12}}(2j_{12}+1)\begin{Bmatrix}j_1 & j_2 & j_{12}\\ j_3 & j_4 & j_{34}\\ j_{13} & j_{24} & j\end{Bmatrix}\begin{Bmatrix}j_1 & j_2 & j_{12}\\ j_{34} & j & j'\end{Bmatrix} = (-)^{2j'}\begin{Bmatrix}j_3 & j_4 & j_{34}\\ j_2 & j' & j_{24}\end{Bmatrix}\begin{Bmatrix}j_{13} & j_{24} & j\\ j' & j_1 & j_3\end{Bmatrix}. \tag{21.231}$$

The $9j$-symbol with one of the j's zero is proportional to a $6j$-symbol:

$$\begin{Bmatrix} j_1 & j_2 & j \\ j_3 & j_4 & j \\ j' & j' & 0 \end{Bmatrix} = \frac{(-)^{j_2+j_3+j+j'}}{\sqrt{(2j+1)(2j'+1)}} \begin{Bmatrix} j_1 & j_2 & j \\ j_4 & j_3 & j' \end{Bmatrix} . \tag{21.232}$$

Algebraic formulas for the the commonly occurring $9j$-symbol:

$$\begin{Bmatrix} \ell & \ell' & L \\ j & j' & J \\ 1/2 & 1/2 & S \end{Bmatrix} , \tag{21.233}$$

for $S = 0, 1$ are given by Matsunobu and Takebe [18]. Values of other special $9j$-symbols can be found in Edmonds [2], Brink and Satchler [17], or Rotenberg, Bivins, Metropolis, and Wooten [16]. The coupling of five and more angular momenta can be done in similar ways as described in this section, but the recoupling coefficients are not used as much in the literature, so we stop here in our discussion of angular momentum coupling.

21.4.3 Rotation of coupled vectors

The relation between eigenvectors of angular momentum for a coupled system described in two coordinate frames Σ and Σ' is given by a rotation operator $U(R)$ for the combined system, $\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2$. Since $\mathbf{J}_1$ and $\mathbf{J}_2$ commute, the rotation operator can be written in two ways:

$$U_{\mathbf{J}}(R) = e^{i\hat{\mathbf{n}}\cdot\mathbf{J}\theta/\hbar} = e^{i\hat{\mathbf{n}}\cdot\mathbf{J}_1\theta/\hbar}\, e^{i\hat{\mathbf{n}}\cdot\mathbf{J}_2\theta/\hbar} = U_{\mathbf{J}_1}(R)\, U_{\mathbf{J}_1}(R) . \tag{21.234}$$

Operating with on (21.189) with $U_{\mathbf{J}}(R)$ and multiplying on the left by the adjoint of Eq. (21.189) gives:

$$\begin{aligned} \langle\, (j_1, j_2)\, j, m \,|\, U_{\mathbf{J}}(R) \,|\, (j_1, j_2)\, j, m' \,\rangle = \sum_{m_1, m_2, m_1', m_2'} & \langle\, j_1, m_1 \,|\, U_{\mathbf{J}_1}(R) \,|\, j_1, m_1' \,\rangle \langle\, j_2, m_2 \,|\, U_{\mathbf{J}_1}(R) \,|\, j_2, m_2' \,\rangle \\ & \times \langle\, (j_1, j_2)\, j, m \,|\, j_1, m_1, j_2, m_2 \,\rangle \langle\, j_1, m_1', j_2, m_2' \,|\, (j_1, j_2)\, j, m' \,\rangle . \end{aligned} \tag{21.235}$$

Here we have used the fact that the matrix elements of the rotation operator is diagonal in the total angular momentum quantum number j. But from Definition 35, matrix elements of the rotation operator are just the D-functions, so (21.235) becomes:

$$\begin{aligned} D^{(j)}_{m,m'}(R) = & \\ \sum_{m_1, m_2, m_1', m_2'} & D^{(j_1)}_{m_1, m_1'}(R)\, D^{(j_2)}_{m_2, m_2'}(R) \langle\, (j_1, j_2)\, j, m \,|\, j_1, m_1, j_2, m_2 \,\rangle \langle\, j_1, m_1', j_2, m_2' \,|\, (j_1, j_2)\, j, m' \,\rangle . \end{aligned} \tag{21.236}$$

Eq. (21.236) is called the **Clebsch-Gordan series**.[11] Another form of it is found by multiplying (21.236) through by another Clebsch-Gordan coefficient and using relations (21.191):

$$\begin{aligned} \sum_m \langle\, j_1, m_1, j_2, m_2 \,|\, (j_1, j_2)\, j, m \,\rangle\, D^{(j)}_{m,m'}(R) & \\ = \sum_{m_1', m_2'} & D^{(j_1)}_{m_1, m_1'}(R)\, D^{(j_2)}_{m_2, m_2'}(R) \langle\, j_1, m_1', j_2, m_2' \,|\, (j_1, j_2)\, j, m' \,\rangle . \end{aligned} \tag{21.237}$$

[11] According to Rotenberg, et. al. [16], A. Clebsch and P. Gordan had little to do with what physicists call the Clebsch-Gordan series.

Exercise 68. Using the infinitesimal expansions:

$$\begin{aligned} D^{(j)}_{m,m'}(\hat{\mathbf{n}}_z,\Delta\theta) &= \delta_{m,m'} + i\,m\,\delta_{m,m'}\,\Delta\theta + \cdots \\ D^{(j)}_{m,m'}(\hat{\mathbf{n}}_\pm,\Delta\theta) &= \delta_{m,m'} + i\,A(j,\mp m')\,\delta_{m,m'\pm1}\,\Delta\theta + \cdots\ , \end{aligned} \tag{21.238}$$

evaluate the Clebsch-Gordan series, Eq. (21.237), for infinitesimal values of θ and for $\hat{\mathbf{n}} = \hat{\mathbf{n}}_z$ and $\hat{\mathbf{n}}_\pm$ to show that Clebsch-Gordan series reproduces Eqs. (21.193) and (21.194). That is, the Clebsch-Gordan series *determines* the Clebsch-Gordan coefficients.

Multiplication of Eq. (21.237) again by a Clebsch-Gordan coefficient and summing over j and m' gives a third relation between D-functions:

$$\begin{aligned} &D^{(j_1)}_{m_1,m_1'}(R)\,D^{(j_2)}_{m_2,m_2'}(R) \\ &\qquad = \sum_{j,m,m'} \langle\, j_1,m_1,j_2,m_2 \,|\, (j_1,j_2)\,j,m\,\rangle\,\langle\, j_1,m_1',j_2,m_2' \,|\, (j_1,j_2)\,j,m'\,\rangle\, D^{(j)}_{m,m'}(R)\,. \end{aligned} \tag{21.239}$$

In terms of $3j$-symbols, (21.239) becomes:

$$D^{(j_1)}_{m_1,m_1'}(R)\,D^{(j_2)}_{m_2,m_2'}(R) = \sum_{j,m,m'} (2j+1) \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j \\ m_1' & m_2' & m' \end{pmatrix} D^{(j)\,*}_{m,m'}(R)\,. \tag{21.240}$$

For integer values of $j_1 = \ell_1$ and $j_2 = \ell_2$ and $m_1 = m_2 = 0$, (21.240) reduces to:

$$C_{\ell_1,m_1}(\Omega)\,C_{\ell_2,m_2}(\Omega) = \sum_{\ell,m} (2\ell+1) \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ m_1 & m_2 & m \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ 0 & 0 & 0 \end{pmatrix} C^*_{\ell,m}(\Omega)\,. \tag{21.241}$$

Using the orthogonality of the spherical harmonics, Eq. (21.241) can be used to find the integral over three spherical harmonics:

$$\int \mathrm{d}\Omega\, C_{\ell_1,m_1}(\Omega)\,C_{\ell_2,m_2}(\Omega)\,C_{\ell_3,m_3}(\Omega) = 4\pi \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}. \tag{21.242}$$

A useful formula can be found from Eq. (21.241) by setting $\ell_1 = 1$ and $m_1 = 0$. Then (21.241) becomes:

$$\begin{aligned} \cos\theta\, C_{\ell_2,m_2}(\theta,\phi) = {}& (2\ell_2+3) \begin{pmatrix} 1 & \ell_2 & \ell_2+1 \\ 0 & m_2 & -m_2 \end{pmatrix} \begin{pmatrix} 1 & \ell_2 & \ell_2+1 \\ 0 & 0 & 0 \end{pmatrix} C^*_{\ell_2+1,-m_2}(\theta,\phi) \\ & + (2\ell_2-1) \begin{pmatrix} 1 & \ell_2 & \ell_2-1 \\ 0 & m_2 & -m_2 \end{pmatrix} \begin{pmatrix} 1 & \ell_2 & \ell_2-1 \\ 0 & 0 & 0 \end{pmatrix} C^*_{\ell_2-1,-m_2}(\theta,\phi) \end{aligned} \tag{21.243}$$

Using Table (21.2) gives the result:

$$\cos\theta\, Y_{\ell,m}(\theta,\phi) = \sqrt{\frac{(\ell+m+1)(\ell-m+1)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1,m}(\theta,\phi) + \sqrt{\frac{(\ell+m)(\ell-m)}{(2\ell-1)(2\ell+1)}}\, Y_{\ell-1,m}(\theta,\phi)\,. \tag{21.244}$$

Similarly, setting $\ell = 1$ and $m = \pm 1$ in (21.241), gives two additional equations:

$$\begin{aligned} \sin\theta\, e^{+i\phi}\, Y_{\ell,m}(\theta,\phi) &= -\sqrt{\frac{(\ell+m+1)(\ell+m+2)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1,m+1}(\theta,\phi) + \sqrt{\frac{(\ell-m)(\ell-m-1)}{(2\ell-1)(2\ell+1)}}\, Y_{\ell-1,m+1}(\theta,\phi) \\ \sin\theta\, e^{-i\phi}\, Y_{\ell,m}(\theta,\phi) &= +\sqrt{\frac{(\ell-m+1)(\ell-m+2)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1,m-1}(\theta,\phi) - \sqrt{\frac{(\ell+m)(\ell+m-1)}{(2\ell-1)(2\ell+1)}}\, Y_{\ell-1,m-1}(\theta,\phi)\,. \end{aligned} \tag{21.245}$$

21.5 Tensor operators

The key to problems involving angular momentum matrix elements of operators is to write the operators in terms of tensor operators, and then use powerful theorems regarding the matrix elements of these tensors. The most important theorem is the Wigner-Eckart theorem, which will be discussed in the next section. Others are discussed in the next section where we also give several examples of the use of these theorems.

21.5.1 Tensor operators and the Wigner-Eckart theorem

Definition 36 (tensor operator)**.** An irreducible tensor operator $T_{k,q}$ of **rank** k and **component** q, with $-k \le q \le +k$, is defined so that under rotation of the coordinate system, it transforms as:

$$U_{\mathbf{J}}(R)\, T_{k,q}\, U^{\dagger}_{\mathbf{J}}(R) = \sum_{q'=-k}^{+k} T_{k,q'}\, D^{(k)}_{q',q}(R)\,. \tag{21.246}$$

where $D^{(k)}_{q,q'}(R)$ is the rotation matrix. The infinitesimal version of (21.246) is:

$$[\, J_i\,, T_{k,q}\,] = \sum_{q'=-k}^{+k} T_{k,q'}\, \langle\, k,q' \,|\, J_i \,|\, k,q\,\rangle\,, \tag{21.247}$$

which gives the equations:

$$[\, J_{\pm}, T_{k,q}\,] = \hbar\, A(k,\mp q)\, T_{k,q\pm1}\,, \qquad [\, J_z, T_{k,q}\,] = \hbar\, q\, T_{k,q}\,. \tag{21.248}$$

Definition 37 (Hermitian tensor operator)**.** The usual definition of a Hermitian tensor operator for integer rank k, and the one we will adopt here, is:

$$T^{\dagger}_{k,q} = (-)^q\, T_{k,-q}\,. \tag{21.249}$$

$R_{1,q}$ and $J_{1,q}$, defined above, and the spherical harmonics satisfies this definition and are Hermitian operators. A second definition, which preserves the Hermitian property for tensor products (see Theorem 45 below) is:

$$T^{\dagger}_{k,q} = (-)^{k-q}\, T_{k,-q}\,. \tag{21.250}$$

The only difference between the two definitions is a factor of i^k.

The adjoint operator $T^{\dagger}_{k,q}$ transforms according to:

$$U(R)\, T^{\dagger}_{k,q}\, U^{\dagger}(R) = \sum_{q'=-k}^{+k} T^{\dagger}_{k,q'}\, D^{(k)\,*}_{q',q}(R) = \sum_{q'=-k}^{+k} T^{\dagger}_{k,q'}\, D^{(k)}_{q,q'}(R^{-1})\,. \tag{21.251}$$

Or putting $R \to R^{-1}$, this can be written as:

$$U^{\dagger}(R)\, T^{\dagger}_{k,q}\, U(R) = \sum_{q'=-k}^{+k} T^{\dagger}_{k,q'}\, D^{(k)}_{q,q'}(R)\,. \tag{21.252}$$

For tensors of half-integer rank, the definition of a Hermitian tensor operator does not work since, for this case, the Hermitian adjoint, taken twice, does not reproduce the same tensor. So a definition of Hermitian is not possible for half-integer operators.

Example 34. The operator made up of the components of the angular momentum operator and defined by:

$$J_{1,q} = \begin{cases} -(J_x + i\,J_y)/\sqrt{2}\,, & \text{for } q = +1, \\ J_z\,, & \text{for } q = 0, \\ +(J_x - i\,J_y)/\sqrt{2}\,, & \text{for } q = -1, \end{cases} \tag{21.253}$$

is a tensor operator of rank one. Since (J_x, J_y, J_z) are Hermitian operators, $J_{1,q}$ satisfies $J^\dagger_{1,q} = (-)^q J_{1,-q}$, and therefore is a Hermitian tensor operator.

Example 35. The spherical harmonics $Y_{k,q}(\Omega)$, considered as operators in coordinate space, are tensor operators. Eqs. (21.25) mean that:

$$[\,J_\pm, Y_{k,q}(\Omega)\,] = \hbar\,A(k, \mp q)\,Y_{k,q\pm1}(\Omega)\,, \qquad [\,J_z, Y_{k,q}(\Omega)\,] = \hbar\,q\,Y_{k,q}(\Omega)\,, \tag{21.254}$$

The reduced spherical harmonics $C_{k,q}(\Omega)$, given in Definition 32, are also tensor operators of rank k component q.

Example 36. The operator $R_{1,q}$ made up of components of the coordinate vector (X, Y, Z) and defined by:

$$R_{1,q} = \begin{cases} -(X + i\,Y)/\sqrt{2}\,, & \text{for } q = +1, \\ Z\,, & \text{for } q = 0, \\ +(X - i\,Y)/\sqrt{2}\,, & \text{for } q = -1. \end{cases} \tag{21.255}$$

where X, Y, and Z are coordinate operators, is a tensor operator of rank one. Using $[\,X_i, L_j\,] = i\hbar\,\epsilon_{ijk}X_k$, one can easily check that Eq. (21.248) is satisfied. Note that since (X, Y, Z) are all Hermitian operators, $R_{1,q}$ satisfies $R^\dagger_{1,q} = (-)^q\,R_{1,-q}$ and so $R_{1,q}$ is a Hermitian tensor operator.

The tensor operator $R_{1,q}$ is a special case of a solid harmonic, defined by:

Definition 38 (solid harmonic)**.** A solid harmonic $R_{k,q}$ is defined by:

$$R_{k,q} = R^k\,C_{k,q}(\Omega)\,. \tag{21.256}$$

Solid harmonics, like the reduced spherical harmonics, are tensor operators of rank k component q.

Example 37. The operator made up of components of the linear momentum and defined by:

$$P_{1,q} = \begin{cases} -(P_x + i\,P_y)/\sqrt{2}\,, & \text{for } q = +1, \\ P_z\,, & \text{for } q = 0, \\ +(P_x - i\,P_y)/\sqrt{2}\,, & \text{for } q = -1. \end{cases} \tag{21.257}$$

where P_x, P_y, and P_z are momentum operators, is a tensor operator of rank one. Using $[\,P_i, L_j\,] = i\hbar\,\epsilon_{ijk}P_k$, one can easily check that Eq. (21.248) is satisfied. Note that since (P_x, P_y, P_z) are all Hermitian operators, $P_{1,q}$ satisfies $P^\dagger_{1,q} = (-)^q\,P_{1,-q}$ and so $P_{1,q}$ is a Hermitian tensor operator.

Finally let us define spherical unit vectors $\hat{\mathbf{e}}_q$ by:

$$\hat{\mathbf{e}}_q = \begin{cases} -(\hat{\mathbf{e}}_x + i\,\hat{\mathbf{e}}_y)/\sqrt{2}\,, & \text{for } q = +1, \\ \hat{\mathbf{e}}_z\,, & \text{for } q = 0, \\ +(\hat{\mathbf{e}}_x - i\,\hat{\mathbf{e}}_y))/\sqrt{2}\,, & \text{for } q = -1. \end{cases} \tag{21.258}$$

These spherical unit vectors are *not* operators. The complex conjugate satisfies: $\hat{\mathbf{e}}^*_q = (-)^q\,\hat{\mathbf{e}}_{-q}$. They also and obey the orthogonality and completeness relations:

$$\hat{\mathbf{e}}_q \cdot \hat{\mathbf{e}}^*_{q'} = \delta_{q,q'}\,, \qquad \sum_q \hat{\mathbf{e}}_q\,\hat{\mathbf{e}}^*_q = \sum_q (-)^q \hat{\mathbf{e}}_q\,\hat{\mathbf{e}}_{-q} = \mathbf{1}\,. \tag{21.259}$$

where $\mathbf{1} = \hat{\mathbf{e}}_x\,\hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y\,\hat{\mathbf{e}}_y + \hat{\mathbf{e}}_z\,\hat{\mathbf{e}}_z$ is the unit dyadic. Any vector operator can be expanded in terms of spherical tensors using these spherical unit vectors. For example, the vector operator $\mathbf{R}$ can be written as:

$$\mathbf{R} = \sum_q (-)^q\, R_{1,q}\,\hat{\mathbf{e}}_{-q}\,, \qquad \text{where} \qquad R_{1,q} = \mathbf{R}\cdot\hat{\mathbf{e}}_q\,. \tag{21.260}$$

Exercise 69 (Edmonds)**.** Let us define a vector operator $\mathbf{S} = \mathbf{e}_x\,S_x + \mathbf{e}_y\,S_y + \mathbf{e}_z\,S_z$, which operates on vectors, by:

$$S_i = i\hbar\,\hat{\mathbf{e}}_i\times\,, \qquad \text{for } i = (x,y,z). \tag{21.261}$$

Show that:

$$S^2\,\hat{\mathbf{e}}_q = \hbar^2\,2\,\hat{\mathbf{e}}_q\,, \qquad S_z\,\hat{\mathbf{e}}_q = \hbar\,q\,\hat{\mathbf{e}}_q\,. \tag{21.262}$$

That is, $\mathbf{S}$ is vector operator for spin one.

Angular momentum matrix elements of irreducible tensor operators with respect to angular momentum eigenvectors are proportional to a Clebsch-Gordan coefficient, or $3j$-symbol, which greatly simplifies calculation of these quantities. The Wigner-Eckart theorem [19, 20], which we now prove, states that fact:

Theorem 44 (Wigner-Eckart)**.** *Angular momentum matrix elements of an irreducible tensor operator $T(k,q)$ is given by:*

$$\begin{aligned}\langle\, j,m\,|\,T_{k,q}\,|\,j',m'\,\rangle &= (-)^{j'-m'}\,\frac{\langle\, j,m,j',-m'\,|\,(j,j')\,k,q\,\rangle}{\sqrt{2k+1}}\,\langle\, j\,\|\,T_k\,\|\,j'\,\rangle\,,\\ &= (-)^{j-m}\begin{pmatrix} j & k & j'\\ -m & q & m'\end{pmatrix}\langle\, j\,\|\,T_k\,\|\,j'\,\rangle\,,\\ &= (-)^{2k}\,\frac{\langle\, j',m',k,q\,|\,(j',k)\,j,m\,\rangle}{\sqrt{2j+1}}\,\langle\, j\,\|\,T_k\,\|\,j'\,\rangle\,.\end{aligned} \tag{21.263}$$

Here $\langle\, j\,\|\,T_k\,\|\,j'\,\rangle$ is called the ***reduced*** *matrix element, and is independent of m, m', and q, which is the whole point of the theorem.*

Proof. Eq. (21.246) can be written as:

$$U(R)\,T_{k,q'}\,U^\dagger(R) = \sum_{q=-k}^{+k} T^\dagger_{k,q}\,D^{(k)}_{q,q'}(R)\,. \tag{21.264}$$

Matrix elements of this equation gives:

$$\begin{aligned}\sum_{q=-k}^{+k}\langle\, j,m\,|\,T_{k,q}\,|\,j',m'\,\rangle\,D^{(j)}_{q,q'}(R) &= \sum_{m'',m'''} D^{(j)}_{m,m''}(R)\,D^{(j)\,*}_{m',m'''}(R)\,\langle\, j,m''\,|\,T_{k,q}\,|\,j',m'''\,\rangle\\ &= \sum_{m'',m'''}(-)^{m'-m'''}\,D^{(j)}_{m,m''}(R)\,D^{(j')}_{-m',-m'''}(R)\,\langle\, j,m''\,|\,T_{k,q}\,|\,j',m'''\,\rangle\end{aligned} \tag{21.265}$$

Now let $m'\to -m'$ and $m'''\to -m'''$, so that (21.265) becomes:

$$\begin{aligned}&\sum_{q=-k}^{+k}\langle\, j,m\,|\,T_{k,q}\,|\,j',-m'\,\rangle\,D^{(j)}_{q,q'}(R)\\ &\qquad = \sum_{m'',m'''}(-)^{m'-m'''}\,D^{(j)}_{m,m''}(R)\,D^{(j')}_{m',m'''}(R)\,\langle\, j,m''\,|\,T_{k,q}\,|\,j',-m'''\,\rangle\end{aligned} \tag{21.266}$$

Comparison with Eq. (21.237) gives:

$$\langle\, j,m\,|\,T_{k,q}\,|\,j',-m'\,\rangle = (-)^{j'+m'}\,\langle\, j,m,j',m'\,|\,(j,j')\,k,q\,\rangle\, f(j,j',k)\,, \tag{21.267}$$

where $f(j,j',k)$ is some function of j, j', and k, and independent of m, m', and q. Choosing $f(j,j',k)$ to be:

$$f(j,j',k) = \frac{\langle\, j\,\|\,T_k\,\|\,j'\,\rangle}{\sqrt{2k+1}}\,, \tag{21.268}$$

proves the theorem as stated. □

Definition 39 (tensor product)**.** Let $T_{k_1,q_1}(1)$ and $T_{k_2,q_2}(2)$ be tensor operators satisfying Definition 36. Then the tensor product of these two operators is defined by:

$$[\,T_{k_1}(1)\otimes T_{k_2}(2)\,]_{k,q} = \sum_{q_1,q_2}\langle\, k_1,q_1,k_2,q_2\,|\,(k_1,k_2)\,k,q\,\rangle\, T_{k_1,q_1}(1)\,T_{k_2,q_2}(2)\,. \tag{21.269}$$

Theorem 45. *The tensor product of Definition 39 is a tensor operator also.*

Proof. The proof relies on the Clebsch-Gordan series, and is left to the reader. □

Theorem 45 means that the Wigner-Eckart theorem applies equally well to tensor products. The Hermitian property for tensor products is preserved if we use the second definition, Eq. (21.250); it is *not* preserved with the usual definition, Eq. (21.249).

Example 38. The tensor product of two commuting vectors

$$[\,R_1(1)\otimes R_1(2)\,]_{k,q} = \sum_{q_1,q_2}\langle\, 1,q_1,1,q_2\,|\,(1,1)\,k,q\,\rangle\, R_{1,q_1}(1)\,R_{1,q_2}(2)\,, \tag{21.270}$$

where $R_{1,q_1}(1)$ and $R_{1,q_2}(2)$ are tensor operators of rank one defined by Eq. (21.255), gives tensor operators of rank $k=0$, 1 and 2. For $k=0$, the tensor product is:

$$\begin{aligned}[\,R_1(1)\otimes R_1(2)\,]_{0,0} &= \sum_{q_1,q_2}\langle\, 1,q_1,1,q_2\,|\,(1,1)\,0,0\,\rangle\, R_{1,q_1}(1)\,R_{1,q_2}(2)\\ &= \frac{-1}{\sqrt{3}}\sum_q (-)^q\, R_{1,q}(1)\,R_{1,-q}(2) = \frac{-1}{\sqrt{3}}\,\mathbf{R}(1)\cdot\mathbf{R}(2)\,,\end{aligned} \tag{21.271}$$

which is a scalar under rotations. For $k=1$, the tensor product is:

$$[\,R_1(1)\otimes R_1(2)\,]_{1,q} = \sum_{q_1,q_2}\langle\, 1,q_1,1,q_2\,|\,(1,1)\,1,q\,\rangle\, R_{1,q_1}(1)\,R_{1,q_2}(2)\,, \tag{21.272}$$

so that using Table 21.1, for $q=+1$, we find:

$$\begin{aligned}[\,R_1(1)\otimes R_1(2)\,]_{1,1} &= \frac{1}{\sqrt{2}}\,(\,R_{1,1}(1)R_{1,0}(2) - R_{1,0}(1)R_{1,1}(2)\,)\\ &= \frac{-i}{2}\,\big(\,(Y(1)Z(2)-Z(1)Y(2)\,) + i\,(Z(1)X(2)-X(2)Z(1)\,)\big) = \frac{i}{\sqrt{2}}\,[\,\mathbf{R}(1)\times\mathbf{R}(2)\,]_{1,1}\,,\end{aligned} \tag{21.273}$$

with similar expressions for $q=0,-1$. So for $q=1,0,-1$, we find:

$$[\,R_1(1)\otimes R_1(2)\,]_{1,q} = \frac{i}{\sqrt{2}}\,[\,\mathbf{R}(1)\times\mathbf{R}(2)\,]_{1,q}\,, \tag{21.274}$$

which is a pseudovector under rotations. For $k = 2$, the five q components are given by,

$$\begin{aligned}
[\, R_1(1) \otimes R_1(2) \,]_{2,\pm 2} &= R_{1,\pm 1}(1)\, R_{1,\pm 1}(2)\,, \\
[\, R_1(1) \otimes R_1(2) \,]_{2,\pm 1} &= \frac{1}{\sqrt{2}} \left(R_{1,\pm 1}(1)\, R_{1,0}(2) + R_{1,0}(1)\, R_{1,\pm 1}(2) \right), \\
[\, R_1(1) \otimes R_1(2) \,]_{2,\ 0} &= \frac{1}{\sqrt{6}} \left(R_{1,1}(1)\, R_{1,-1}(2) + 2\, R_{1,0}(1)\, R_{1,0}(2) + R_{1,1}(-1)\, R_{1,1}(2) \right),
\end{aligned} \tag{21.275}$$

which can be written in terms of the Cartesian components of the traceless symmetric tensor:

$$R_{ij}(1,2) = \frac{1}{2} \left(R_i(1) R_j(2) + R_j(1) R_i(2) \right) - \frac{1}{3}\, \delta_{ij} \left(\mathbf{R}(1) \cdot \mathbf{R}(2) \right). \tag{21.276}$$

Definition 40 (Scalar product)**.** For the zero rank tensor product of two rank one tensors, it is useful to have a special definition, called the **scalar product**, so that it agrees with the usual dot product of vectors. So we define:

$$\begin{aligned}
[\, T_k(1) \odot T_k(2) \,] &= \sum_q (-)^q\, T_{k,q}(1)\, T_{k,-q}(2) = \sum_q T_{k,q}(1)\, T^\dagger_{k,q}(2) = \sum_q T^\dagger_{k,q}(1)\, T_{k,q}(2) \\
&= \sqrt{2k+1}\, (-)^k\, [\, T_k(1) \otimes T_k(2) \,]_{0,0}\,.
\end{aligned} \tag{21.277}$$

Example 39. The scalar product of two vectors is just the vector dot product. We find:

$$[\, R_1(1) \odot R_1(2) \,] = \sum_q (-)^q\, R_{1,q}(1)\, R_{1,-q}(2) = \mathbf{R}(1) \cdot \mathbf{R}(2)\,. \tag{21.278}$$

Example 40. An important example of a scalar product is given by writing the addition theorem for spherical harmonics, Eq. (21.183), as a tensor product:

$$P_k(\cos\gamma) = \frac{4\pi}{2\ell+1} \sum_{q=-k}^{+k} Y_{k,q}(\Omega)\, Y^*_{k,q}(\Omega') = \sum_{q=-k}^{+k} C_{k,q}(\Omega)\, C^*_{k,q}(\Omega') = [\, C_k(\Omega) \odot C_k(\Omega') \,]\,. \tag{21.279}$$

21.5.2 Reduced matrix elements

The Wigner-Eckart theorem enables us to calculate matrix elements of tensor operators for different values of (m, m', q) if we know the reduced matrix element, so it is useful to have a table of reduced matrix elements for a number of operators that might enter into calculations. We can generate this table and find reduced matrix elements by computing only *one* matrix element for certain specified values of (m, m', q). For integer rank tensors, we often just take the case when $m = m' = q = 0$. Then the reduced matrix element is found by using the Wigner-Eckart theorem backwards.

unit operator

Reduced matrix elements of the unit operator are easily found to be:

$$\langle\, j \,\|\, 1 \,\|\, j' \,\rangle = \delta_{j,j'}\, \sqrt{2j+1}\,. \tag{21.280}$$

angular momentum

Reduced matrix elements for the angular momentum tensor operator J are easily found by noting that $\langle\, j, m \,|\, J_{1,0} \,|\, j', m' \,\rangle = m\, \delta_{j,j'} \delta_{m,m'}$, and using Table 21.2 for the appropriate $3j$-symbol. This gives:

$$\langle\, j \,\|\, J \,\|\, j' \,\rangle = \hbar\, \delta_{j,j'}\, \sqrt{2j(2j+1)(2j+2)}/2\,. \tag{21.281}$$

and

$$\langle\, \ell \,\|\, L \,\|\, \ell' \,\rangle = \hbar\, \delta_{\ell,\ell'}\, \sqrt{\ell(\ell+1)(2\ell+1)}\,. \tag{21.282}$$

A special case is:

$$\langle\, 1/2 \,\|\, \sigma \,\|\, 1/2 \,\rangle = \sqrt{6}\,. \tag{21.283}$$

Exercise 70. The angular momentum tensor operator $J_{1,q}$ can be written as:

$$J_{1,q} = \begin{cases} -J_+/\sqrt{2}\,, & \text{for } q = +1, \\ J_z\,, & \text{for } q = \;\; 0, \\ +J_-/\sqrt{2}\,, & \text{for } q = -1. \end{cases} \tag{21.284}$$

Compute directly $J_{1,q}\,|\,j,m\,\rangle$ for all values of q. Now use the reduced matrix element given in Eq. (21.281) and the Wigner-Eckart theorem to compute:

$$J_{1,q}\,|\,j,m\,\rangle = \sum_{j',m'} |\,j',m'\,\rangle\,\langle\,j',m'\,|\,J_{1,q}\,|\,j,m\,\rangle\,, \tag{21.285}$$

and show that you get the same result.

spherical harmonics

The reduced matrix elements of spherical harmonics can be found most easily in coordinate space using Eq. (21.242). This gives:

$$\langle\, \ell \,\|\, C_k \,\|\, \ell' \,\rangle = (-)^\ell \sqrt{(2\ell+1)(2\ell'+1)} \begin{pmatrix} \ell & k & \ell' \\ 0 & 0 & 0 \end{pmatrix}. \tag{21.286}$$

and

$$\langle\, \ell \,\|\, Y_k \,\|\, \ell' \,\rangle = (-)^\ell \sqrt{\frac{(2\ell+1)(2\ell'+1)(2k+1)}{4\pi}} \begin{pmatrix} \ell & k & \ell' \\ 0 & 0 & 0 \end{pmatrix}. \tag{21.287}$$

The reduced matrix elements of the solid harmonics involve radial integrals, which we have ignored up to now. Adding radial quantum numbers to the matrix elements gives:

$$\langle\, n,\ell \,\|\, R_k \,\|\, n',\ell' \,\rangle = (-)^\ell \sqrt{(2\ell+1)(2\ell'+1)} \begin{pmatrix} \ell & k & \ell' \\ 0 & 0 & 0 \end{pmatrix} \int_0^\infty r^2\,dr\, R_{n,\ell}(r)\, r^k\, R_{n',\ell'}(r)\,, \tag{21.288}$$

where $R_{n,\ell}(r)$ are (real) radial wave functions for the state (n,ℓ), normalized to one with measure $\mu = r^2$.

linear momentum (the gradient formula)

The reduced matrix elements of the momentum operator require expansion of the gradient operator in spherical coordinates. Let us first note that from Eqs. (21.19), (21.20), and (21.23), in a coordinate basis:

$$\begin{aligned} P_z &= \frac{\hbar}{i}\frac{\partial}{\partial z} = \frac{\hbar}{i}\left\{ \cos\theta\,\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\,\frac{\partial}{\partial\theta} \right\} \\ &= \frac{\hbar}{i}\,\cos\theta\,\frac{\partial}{\partial r} + \frac{i\,\sin\theta}{2\,r}\left[\, e^{-i\phi}\,L_+ - e^{+i\phi}\,L_- \right]. \end{aligned} \tag{21.289}$$

Using Eqs. 21.245, we find:

$$\begin{aligned} &\sin\theta\, e^{-i\phi}\, L_+\, Y_{\ell,m}(\theta,\phi) = \hbar\, A(\ell,-m)\, \sin\theta\, e^{-i\phi}\, Y_{\ell,m+1}(\theta,\phi) \\ &= \hbar\sqrt{(\ell-m)(\ell+m+1)}\left\{ \sqrt{\frac{(\ell-m)(\ell-m+1)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1,m}(\theta,\phi) - \sqrt{\frac{(\ell+m+1)(\ell+m)}{(2\ell-1)(2\ell+1)}}\, Y_{\ell-1,m}(\theta,\phi) \right\} \\ &= \hbar\left\{ (\ell-m)\sqrt{\frac{(\ell+m+1)(\ell-m+1)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1,m}(\theta,\phi) - (\ell+m+1)\sqrt{\frac{(\ell+m)(\ell-m)}{(2\ell-1)(2\ell+1)}}\, Y_{\ell-1,m}(\theta,\phi) \right\}, \end{aligned} \tag{21.290}$$

and

$$\sin\theta\, e^{+i\phi}\, L_-\, Y_{\ell,m}(\theta,\phi) = \hbar\, A(\ell,m)\, \sin\theta\, e^{+i\phi}\, Y_{\ell,m-1}(\theta,\phi)$$
$$= \hbar\sqrt{(\ell+m)(\ell-m+1)}\left\{ -\sqrt{\frac{(\ell+m)(\ell+m+1)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1,m}(\theta,\phi) + \sqrt{\frac{(\ell-m+1)(\ell-m)}{(2\ell-1)(2\ell+1)}}\, Y_{\ell-1,m}(\theta,\phi)\right\}$$
$$= \hbar\left\{ -(\ell+m)\sqrt{\frac{(\ell+m+1)(\ell-m+1)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1,m}(\theta,\phi) + (\ell+m+1)\sqrt{\frac{(\ell-m)(\ell+m)}{(2\ell-1)(2\ell+1)}}\, Y_{\ell-1,m}(\theta,\phi)\right\}. \tag{21.291}$$

So

$$\frac{i}{2}\left[\sin\theta\, e^{-i\phi}\, L_+ - \sin\theta\, e^{+i\phi}\, L_-\right] Y_{\ell,m}(\theta,\phi)$$
$$= -\frac{\hbar}{i}\left\{ \ell\sqrt{\frac{(\ell+m+1)(\ell-m+1)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1,m}(\theta,\phi) - (\ell+1)\sqrt{\frac{(\ell+m)(\ell-m)}{(2\ell-1)(2\ell+1)}}\, Y_{\ell-1,m}(\theta,\phi)\right\}. \tag{21.292}$$

So Eq. (21.289) becomes:

$$P_z\, Y_{\ell,m}(\theta,\phi) = \frac{\hbar}{i}\cos\theta\, Y_{\ell,m}(\theta,\phi)\frac{\partial}{\partial r} + \frac{i\sin\theta}{2\,r}\left[e^{-i\phi}\, L_+ - e^{+i\phi}\, L_-\right] Y_{\ell,m}(\theta,\phi)$$
$$= \frac{\hbar}{i}\left[\frac{\partial}{\partial r} - \frac{\ell}{r}\right]\sqrt{\frac{(\ell+m+1)(\ell-m+1)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1,m}(\theta,\phi) + \frac{\hbar}{i}\left[\frac{\partial}{\partial r} + \frac{\ell+1}{r}\right]\sqrt{\frac{(\ell+m)(\ell-m)}{(2\ell-1)(2\ell+1)}}\, Y_{\ell-1,m}(\theta,\phi)\,. \tag{21.293}$$

where we have used Eq. (21.244). So taking angular matrix elements of the above and setting $m = 0$ gives the result:

$$\langle\, \ell', 0\,|\, P_z\,|\, \ell, 0\,\rangle = \begin{cases} \dfrac{\ell+1}{\sqrt{(2\ell+1)(2\ell+3)}}\,\dfrac{\hbar}{i}\left[\dfrac{\partial}{\partial r} - \dfrac{\ell}{r}\right], & \text{for } \ell' = \ell+1, \\ \dfrac{\ell}{\sqrt{(2\ell-1)(2\ell+1)}}\,\dfrac{\hbar}{i}\left[\dfrac{\partial}{\partial r} + \dfrac{\ell+1}{r}\right], & \text{for } \ell' = \ell-1. \end{cases} \tag{21.294}$$

But from the Wigner-Eckart theorem,

$$\langle\, \ell', 0\,|\, P_0\,|\, \ell, 0\,\rangle = (-)^{\ell'}\begin{pmatrix} \ell' & 1 & \ell \\ 0 & 0 & 0 \end{pmatrix}\langle\, \ell'\,\|\, P\,\|\, \ell\,\rangle\,, \tag{21.295}$$

and using the Table 21.2 for the $3j$-symbol, we find:

$$\begin{pmatrix} \ell' & 1 & \ell \\ 0 & 0 & 0 \end{pmatrix} = \begin{cases} (-)^{\ell+1}\,\dfrac{\ell+1}{\sqrt{(\ell+1)(2\ell+1)(2\ell+3)}}, & \text{for } \ell' = \ell+1, \\ (-)^{\ell}\,\dfrac{\ell}{\sqrt{\ell(2\ell-1)(2\ell+1)}}, & \text{for } \ell' = \ell-1. \end{cases} \tag{21.296}$$

So the angular reduced matrix elements of the momentum operator is given by:

$$\langle\, \ell'\,\|\, P\,\|\, \ell\,\rangle = \begin{cases} \sqrt{\ell+1}\,\dfrac{\hbar}{i}\left[\dfrac{\partial}{\partial r} - \dfrac{\ell}{r}\right], & \text{for } \ell' = \ell+1, \\ -\sqrt{\ell}\,\dfrac{\hbar}{i}\left[\dfrac{\partial}{\partial r} + \dfrac{\ell+1}{r}\right], & \text{for } \ell' = \ell-1. \end{cases} \tag{21.297}$$

Remark 38. There exists a simple derivation of Eq. (21.297) due to C. L. Schwartz [16][quoted by Rotenberg, et. al.]. It rests on the observation that:

$$P_i = \frac{1}{2i\hbar}\,[\,X_i, P^2\,]\,, \tag{21.298}$$

and the result obtained in Section 21.1.3 that $P^2 = P_r^2 + L^2/R^2$. Substituting this into (21.298) gives:

$$P_i = \frac{1}{2i\hbar}\,[\,X_i, P_r^2\,] + \frac{1}{2i\hbar}\,[\,X_i, L^2/R^2\,]\,. \tag{21.299}$$

But we find:

$$[\,X_i, P_r^2\,] = P_r\,[\,X_i, P_r\,] + [\,X_i, P_r\,]\,P_r = i\hbar\left\{\,P_r\,\frac{X_i}{R} + \frac{X_i}{R}\,P_r\,\right\} = 2i\hbar\,\frac{X_i}{R}\,P_r\,, \tag{21.300}$$

since $[\,P_r, X_i/R\,] = 0$. So (21.299) becomes:

$$P_i = \frac{X_i}{R}\,P_r + \frac{1}{2i\hbar}\,[\,X_i, L^2/R^2\,]\,. \tag{21.301}$$

The same formula works for spherical tensor operators $P_i \mapsto P_{1,q}$ and $X_i \mapsto R_{1,q}$. Noting that $R_{1,q} = R\,C_{1,q}(\Omega)$, Eq. (21.301) becomes the operator equation:

$$P_{1,q} = C_{1,q}(\Omega)\,P_r + \frac{1}{2i\hbar}\,[\,C_{1,q}(\Omega), L^2/R\,]\,. \tag{21.302}$$

The angular momentum reduced matrix elements of (21.302) gives, in a radial coordinate representation:

$$\langle\,\ell'\,\|\,P\,\|\,\ell\,\rangle = \langle\,\ell'\,\|\,C_1\,\|\,\ell\,\rangle\,\frac{\hbar}{i}\left[\,\frac{\partial}{\partial r} + \frac{1}{r} + \frac{\ell(\ell+1) - \ell'(\ell'+1)}{2r}\,\right] \tag{21.303}$$

Substituting the result for $\langle\,\ell'\,\|\,C_1\,\|\,\ell\,\rangle$ from Eq. (21.286) yields (21.297).

21.5.3 Angular momentum matrix elements of tensor operators

In this section, we give several theorems regarding angular momentum matrix elements of tensor operators. These theorem are the basis for calculating all matrix elements in coupled schemes. The theorems are from Edmonds [2][Chapter 7].

Theorem 46. *Let $T_{k_1,q_1}(1)$ and $S_{k_2,q_2}(1)$ be two tensor operators of rank k_1 and k_2 which act on the* same *angular momentum system, with $[\,T_{k_1,q_1}(1), S_{k_2,q_2}(1)\,] = 0$. Then*

$$\begin{aligned}\langle\,j\,\|\,[\,T_{k_1}(1)\otimes S_{k_2}(1)\,]_k\,\|\,j'\,\rangle \\ = \sqrt{2k+1}\,(-)^{k+j+j'}\sum_{j''}\begin{Bmatrix} k_1 & k_2 & k \\ j' & j & j'' \end{Bmatrix}\langle\,j\,\|\,T_{k_1}(1)\,\|\,j''\,\rangle\,\langle\,j''\,\|\,S_{k_2}(1)\,\|\,j'\,\rangle\,.\end{aligned} \tag{21.304}$$

Proof. Inverting the Wigner-Eckart theorem, Eq. (21.263), using the orthogonality relations, Eq. (21.210)

for $3j$-symbols, and introducing a complete set of states, we have:

$$\begin{aligned}
&\langle\, j \,\|\, [\, T_{k_1}(1) \otimes S_{k_2}(1)\,]_k \,\|\, j' \,\rangle \\
&\qquad = \sum_{m,m',q} (-)^{j-m} \begin{pmatrix} j & k & j' \\ -m & q & m' \end{pmatrix} \langle\, j,m \,|\, [\, T_{k_1}(1) \otimes S_{k_2}(1)\,]_{k,q} \,|\, j',m' \,\rangle \\
&\qquad = \sqrt{2k+1} \sum_{j''} \langle\, j \,\|\, T_{k_1}(1) \,\|\, j'' \,\rangle \langle\, j'' \,\|\, S_{k_2}(1) \,\|\, j' \,\rangle \\
&\times \sum_{\substack{m,m',m'' \\ q_1,q_2,q}} (-)^{k_1-k_2+q+j''-m''} \begin{pmatrix} j & k & j' \\ -m & q & m' \end{pmatrix} \begin{pmatrix} k_1 & k_2 & k \\ q_1 & q_2 & -q \end{pmatrix} \begin{pmatrix} j & k_1 & j'' \\ -m & q_1 & m'' \end{pmatrix} \begin{pmatrix} j'' & k_2 & j' \\ -m'' & q_2 & m' \end{pmatrix} \\
&\qquad = \sqrt{2k+1}\,(-)^{k+j+j'} \sum_{j''} \langle\, j \,\|\, T_{k_1}(1) \,\|\, j'' \,\rangle \langle\, j'' \,\|\, S_{k_2}(1) \,\|\, j' \,\rangle \\
&\quad \times \sum_{\substack{m,m',m'' \\ q_1,q_2,q}} (-)^{p} \begin{pmatrix} k_1 & k_2 & k \\ q_1 & q_2 & q \end{pmatrix} \begin{pmatrix} k_1 & j & j'' \\ -q_1 & m & -m'' \end{pmatrix} \begin{pmatrix} j' & k_2 & j'' \\ -m' & -q_2 & m'' \end{pmatrix} \begin{pmatrix} j' & j & k \\ m' & -m & -q \end{pmatrix} \qquad (21.305)
\end{aligned}$$

where $p = k_1 + k_2 + k + j + j' + j'' + q_1 + q_2 + q + m + m' + m''$ (recall that all k's and q's are integers). This last line is just the $6j$-symbol defined in Eq. (21.219), which completes the proof. Note that if other quantum numbers are needed to complete the states, they should be added to the intermediate sum. □

Theorem 47. *Let $T_{k_1,q_1}(1)$ and $T_{k_2,q_2}(2)$ be two tensor operators which act on parts one and two of a combined system, with $[\, T_{k_1,q_1}(1), T_{k_2,q_2}(2)\,] = 0$. Then*

$$\begin{aligned}
&\langle\, (j_1,j_2)\, j \,\|\, [\, T_{k_1}(1) \otimes S_{k_2}(1)\,]_k \,\|\, (j_1',j_2')\, j' \,\rangle \\
&\qquad = \sqrt{(2k+1)(2j+1)(2j'+1)} \begin{Bmatrix} j_1 & j_1' & k_1 \\ j_2 & j_2' & k_2 \\ j & j' & k \end{Bmatrix} \langle\, j_1 \,\|\, T_{k_1}(1) \,\|\, j_1' \,\rangle \langle\, j_2 \,\|\, T_{k_2}(2) \,\|\, j_2' \,\rangle\,. \qquad (21.306)
\end{aligned}$$

Proof. Here again we invert the Wigner-Echart theorem, and uncouple the states, this time obtaining:

$$\begin{aligned}
&\langle\, (j_1,j_2)\, j \,\|\, [\, T_{k_1}(1) \otimes S_{k_2}(1)\,]_k \,\|\, (j_1',j_2')\, j' \,\rangle \\
&\quad = \sqrt{2k+1} \sum_{m,m'} (-)^{j'+m'} \langle\, j,m,j',m' \,|\, (j,j')\, k,q \,\rangle \langle\, (j_1,j_2)\, j,m \,|\, [\, T_{k_1}(1) \otimes S_{k_2}(1)\,]_{k,q} \,|\, (j_1',j_2')\, j',-m' \,\rangle \\
&\qquad = \sqrt{2k+1} \sum_{\substack{m,m',q_1,q_2 \\ m_1,m_2,m_1',m_2'}} (-)^{j'+m'} \langle\, j,m,j',m' \,|\, (j,j')\, k,q \,\rangle \langle\, k_1,q_1,k_2,q_2 \,|\, (k_1,k_2)\, k,q \,\rangle \\
&\qquad\qquad \times \langle\, j_1,m_1,j_2,m_2 \,|\, (j_1,j_2)\, j,m \,\rangle \langle\, j_1',m_1',j_2',m_2' \,|\, (j_1',j_2')\, j',-m' \,\rangle \\
&\qquad\qquad \times \langle\, j_1,m_1 \,|\, T_{k_1,m_1}(1) \,|\, j_1',m_1' \,\rangle \langle\, j_2,m_2 \,|\, T_{k_2,m_2}(1) \,|\, j_2',m_2' \,\rangle \\
&\qquad\qquad = \sqrt{\frac{2k+1}{(2k_1+1)(2k_2+1)}}\, \langle\, j_1 \,\|\, T_{k_1}(1) \,\|\, j_1' \,\rangle \langle\, j_2 \,\|\, T_{k_2}(2) \,\|\, j_2' \,\rangle \\
&\quad \times \sum_{\substack{m,m',q_1,q_2 \\ m_1,m_2,m_1',m_2'}} (-)^{j_1'+j_2'-j-m_1'-m_2'-m'} \langle\, j,m,j',m' \,|\, (j,j')\, k,q \,\rangle \langle\, k_1,q_1,k_2,q_2 \,|\, (k_1,k_2)\, k,q \,\rangle \\
&\qquad\qquad \times \langle\, j_1,m_1,j_2,m_2 \,|\, (j_1,j_2)\, j,m \,\rangle \langle\, j_1',m_1',j_2',m_2' \,|\, (j_1',j_2')\, j',-m' \,\rangle \\
&\qquad\qquad\qquad \times \langle\, j_1,m_1,j_1',-m_1' \,|\, (j_1,j_1')\, k_1,q_1 \,\rangle \langle\, j_2,m_2,j_2',-m_2' \,|\, (j_2,j_2')\, k_2,q_2 \,\rangle\,. \qquad (21.307)
\end{aligned}$$

Now $m' + m'_1 + m'_2 = 0$, and setting $m'_1 \to -m'_1$ and $m'_2 \to -m'_2$, and noting that

$$\langle\, j'_1, -m'_1, j'_2, -m'_2 \,|\, (j'_1, j'_2)\, j', -m' \,\rangle = (-)^{j'_1+j'_2-j'} \,\langle\, j'_1, m'_1, j'_2, m'_2 \,|\, (j'_1, j'_2)\, j', m' \,\rangle\,, \tag{21.308}$$

the last sum in (21.307) is just the recoupling coefficient:

$$\langle\, (j_1, j_2)\, j\, (j'_1, j'_2)\, j', k, q \,|\, (k_1, k_2)\, k\, (k'_1, k'_2)\, k', k, q \,\rangle$$
$$= \sqrt{(2j+1)(2j'+1)(2k+1)(2k'+1)} \begin{Bmatrix} j_1 & j'_1 & k_1 \\ j_2 & j'_2 & k_2 \\ j & j' & k \end{Bmatrix}, \tag{21.309}$$

and is independent of q. Substitution of (21.309) into (21.307) proves the theorem. □

Theorem 48. *Matrix elements of the scalar product of two tensor operators $T_{k,q}(1)$ and $T_{k,q}(2)$ which act on parts one and two of a coupled system is given by:*

$$\langle\, (j_1, j_2)\, j, m \,|\, [\, T_k(1) \odot T_k(2)\,] \,|\, (j'_1, j'_2)\, j', m' \,\rangle$$
$$= \delta_{j,j'}\, \delta_{m,m'}\, (-)^{j'_1+j_2+j} \begin{Bmatrix} j & j_2 & j_1 \\ k & j'_1 & j'_2 \end{Bmatrix} \langle\, j_1 \,\|\, T_k(1) \,\|\, j'_1 \,\rangle\, \langle\, j_2 \,\|\, T_k(2) \,\|\, j'_2 \,\rangle\,. \tag{21.310}$$

Proof. Here we use definition 40 of the scalar product, set $k = 0$, and put $k_1 = k_2 \to k$ in Theorem 47, and use Eq. (21.232) for the $9j$-symbol with one zero entry. The result follows easily. □

Theorem 49. *The reduced matrix element of a tensor operators $T_{k,q}(1)$ which acts only on part one of a coupled system is given by:*

$$\langle\, (j_1, j_2)\, j \,\|\, T_k(1) \,\|\, (j'_1, j'_2)\, j' \,\rangle$$
$$= \delta_{j_2,j'_2} (-)^{j_1+j_2+j'+k} \sqrt{(2j+1)(2j'+1)} \begin{Bmatrix} j_1 & j & j'_2 \\ j' & j'_1 & k \end{Bmatrix} \langle\, j_1 \,\|\, T_k(1) \,\|\, j'_1 \,\rangle\,. \tag{21.311}$$

Proof. Here we put $T_{0,0}(2) = 1$, and put $k_2 = 0$ and $k = k_1$ in Theorem 47. Using (21.232) again, the result follows. □

Theorem 50. *The reduced matrix element of a tensor operators $T_2(k, q)$ which acts only on part two of a coupled system is given by:*

$$\langle\, (j_1, j_2)\, j \,\|\, T_k(2) \,\|\, (j'_1, j'_2)\, j' \,\rangle$$
$$= \delta_{j_2,j'_2} (-)^{j'_1+j_2+j+k} \sqrt{(2j+1)(2j'+1)} \begin{Bmatrix} j_2 & j & j'_1 \\ j' & j'_2 & k \end{Bmatrix} \langle\, j_2 \,\|\, T_k(2) \,\|\, j'_2 \,\rangle\,. \tag{21.312}$$

Proof. For this case, we put $T_{0,0}(1) = 1$, and put $k_1 = 0$ and $k = k_2$ in Theorem 47 again, from which the result follows. □

Theorem 51 (Projection Theorem)**.** *If $\mathbf{V}$ is a vector operator such that $[\, V_i, J_j \,] = 0$, then*

$$\langle\, j, m \,|\, \mathbf{V} \,|\, j, m' \,\rangle = \langle\, j, m \,|\, (\, \mathbf{V} \cdot \mathbf{J}\,)\, \mathbf{J}/J^2 \,|\, j, m' \,\rangle\,. \tag{21.313}$$

Note that this theorem is valid only for the case when $j' = j$.

Proof. We first note that from the Wigner-Eckart theorem, the left-hand-side of Eq. (21.313) is given by:

$$\langle\, j, m \,|\, V_{1,q} \,|\, j, m' \,\rangle = (-)^{j-m'} \langle\, j, m, j, -m' \,|\, (j, j)1, q \,\rangle\, \langle\, j \,\|\, V \,\|\, j' \,\rangle / \sqrt{3}\,, \tag{21.314}$$

whereas on the right-hand-side of (21.313), we have:

$$\langle\, j,m\,|\,(\,\mathbf{V}\cdot\mathbf{J}\,)\,J_{1,q}/J^2\,|\,j',m'\,\rangle = \sum_{j''m''}\langle\, j,m\,|\,(\,\mathbf{V}\cdot\mathbf{J}\,)\,|\,j'',m''\,\rangle\,\langle\, j'',m''\,|\,J_{1,q}\,|\,j',m'\,\rangle/j(j+1)\,. \tag{21.315}$$

Now

$$\begin{aligned}\langle\, j'',m''\,|\,J_{1,q}\,|\,j',m'\,\rangle &= (-)^{j''-m''}\langle\, j'',m'',j',-m'\,|\,(j,j')1,q\,\rangle\,\langle\, j''\,\|\,J\,\|\,j'\,\rangle/\sqrt{3}\\ &= \delta_{j'',j'}(-)^{j'-m''}\langle\, j',m'',j',-m'\,|\,(j,j')1,q\,\rangle\,\sqrt{j(j+1)(2j+1)}/\sqrt{3}\,,\end{aligned} \tag{21.316}$$

and since $(\,\mathbf{V}\cdot\mathbf{J}\,) = -\sqrt{3}\,T_{0,0}$, where $T_{0,0}$ is a tensor operator of rank zero, again using the Wigner-Eckart theorem,

$$\langle\, j,m\,|\,(\,\mathbf{V}\cdot\mathbf{J}\,)\,|\,j'',m''\,\rangle = \delta_{j,j''}\delta_{m,m''}\,\langle\, j\,\|\,(\,\mathbf{V}\cdot\mathbf{J}\,)\,\|\,j\,\rangle/\sqrt{2j+1}\,. \tag{21.317}$$

Now since $[\,V_i,J_j\,]=0$, Theorem 46 applies with $k=0$ and $k_1=k_2=1$, and using the first line in Table 21.2 and Eq. (21.281), we easily find:

$$\langle\, j\,\|\,(\,\mathbf{V}\cdot\mathbf{J}\,)\,\|\,j\,\rangle = \sqrt{j(j+1)}\,\langle\, j\,\|\,V\,\|\,j\,\rangle\,. \tag{21.318}$$

So combining Eqs. (21.315), (21.316), (21.317), and (21.318), we find that matrix elements of the right-hand-side gives:

$$\langle\, j,m\,|\,(\,\mathbf{V}\cdot\mathbf{J}\,)\,J_{1,q}/J^2\,|\,j',m'\,\rangle = \delta_{j,j'}\,(-)^{j-m'}\langle\, j,m,j,-m'\,|\,(j,j)1,q\,\rangle\,\langle\, j\,\|\,V\,\|\,j\,\rangle/\sqrt{3}\,, \tag{21.319}$$

which is the same result as (21.314), provided that $j=j'$, which proves the theorem. □

Remark 39. The projection theorem is often quoted without proof in elementary texts, and "justified" by stating that the *average* value of the vector $\mathbf{V}$ in good eigenstates of J^2 is given by the projection of $\mathbf{V}$ onto $\mathbf{J}$ in the direction of the angular momentum $\mathbf{J}$. This argument assumes that the average value of the component of $\mathbf{V}$ *perpendicular* to $\mathbf{J}$ vanishes. Usually, one thinks about this as a time average, but time has nothing to do with it! The projection theorem is often used to find the Landé g_J-factor for computing the weak field Zeeman effect in atoms. Of course, one can just apply the appropriate theorems in this section to find matrix elements of any vector operator, the projection theorem is not really needed, but it provides a quick way to get the same result.

21.6 Selected problems in atomic and nuclear physics

In this section, we give several examples of the use of tensor operators in atomic and nuclear physics.

21.6.1 Spin-orbit force in hydrogen

The spin-orbit force for the electron in a hydrogen atom in atomic units is given by a Hamiltonian of the form (see Section 22.3.7):

$$H_{so} = V(R)\,(\,\mathbf{L}\cdot\mathbf{S}\,)/\hbar^2\,. \tag{21.320}$$

Of course it is easy to calculate this in perturbation theory for the states $|\,n,(\ell,s)\,j,m_j\,\rangle$. Since $\mathbf{J}=\mathbf{L}+\mathbf{S}$, and squaring this expression, we find that we can write:

$$\mathbf{L}\cdot\mathbf{S} = \frac{1}{2}\,(\,J^2-L^2-S^2\,)\,, \tag{21.321}$$

so that we find:

$$\langle\,(\ell,s)\,j,m_j\,|\,\mathbf{L}\cdot\mathbf{S}\,|\,(\ell,s)\,j,m_j\,\rangle/\hbar^2 = \frac{1}{2}\big(\,j(j+1)-\ell(\ell+1)-3/4\,\big)\,. \tag{21.322}$$

Since $\mathbf{L}\cdot\mathbf{S} = [\,L\odot S\,]$, we can also find matrix elements of the spin-orbit force using Theorem 48. This gives:

$$\langle\,(\ell,s)\,j,m_j\,|\,[\,L\odot S\,]\,|\,(\ell,s)\,j,m_j\,\rangle/\hbar^2 = (-)^{j+\ell+s}\begin{Bmatrix} j & s & \ell \\ 1 & \ell & s\end{Bmatrix}\langle\,\ell\,\|\,L\,\|\,\ell\,\rangle\,\langle\,s\,\|\,S\,\|\,s\,\rangle/\hbar^2\,. \tag{21.323}$$

Now using the $6j$-tables in Edmonds, we find:

$$\begin{aligned}
\begin{Bmatrix} j & s & \ell \\ 1 & \ell & s\end{Bmatrix} &= (-)^{j+\ell+s}\,\frac{2\,[\,j(j+1)-\ell(\ell+1)-s(s+1)\,]}{\sqrt{2\ell(2\ell+1)(2\ell+2)2s(2s+1)(2s+2)}}\,,\\
\langle\,\ell\,\|\,L\,\|\,\ell\,\rangle/\hbar &= \sqrt{2\ell(2\ell+1)(2\ell+2)}/2\,,\\
\langle\,s\,\|\,S\,\|\,s\,\rangle/\hbar &= \sqrt{2s(2s+1)(2s+2)}/2\,,
\end{aligned} \tag{21.324}$$

so (21.323) becomes simply:

$$\langle\,(\ell,s)\,j,m_j\,|\,[\,L\odot S\,]\,|\,(\ell,s)\,j,m_j\,\rangle/\hbar^2 = \frac{1}{2}\big(\,j(j+1)-\ell(\ell+1)-3/4\,\big)\,. \tag{21.325}$$

in agreement with Eq. (22.163). Of course, using the fancy angular momentum theorems for tensor operators in this case is over-kill! Our point was to show that the theorems give the same result as the simple way. We will find in later examples that the *only* way to do the problem is to use the fancy theorems.

21.6.2 Transition rates for photon emission in Hydrogen

Omit?

21.6.3 Hyperfine splitting in Hydrogen

In this section, we show how to compute the hyperfine energy splitting in hydrogen due to the interaction between the magnetic moment of the proton and the electron. We derive the forces responsible for the splitting in Section 22.3.8 where, in atomic units, we found the Hamiltonian:

$$\bar{H}_{\rm hf} = 2\,\lambda_{\rm p}\left(\frac{m}{M}\right)\alpha^2\,\frac{\mathbf{K}_{\rm e}\cdot\mathbf{S}_{\rm p}/\hbar^2}{\bar{R}^3}\,,\qquad \mathbf{K}_{\rm e} = \mathbf{L}_{\rm e}-\mathbf{S}_{\rm e}+3\,(\mathbf{S}_{\rm e}\cdot\hat{\mathbf{R}})\,\hat{\mathbf{R}}\,, \tag{21.326}$$

where $\mathbf{L}_{\rm e}$ and $\mathbf{S}_{\rm e}$ are the angular momentum and spin operators for the electron, $\mathbf{S}_{\rm p}$ is the spin operator for the proton, and $\hat{\mathbf{R}}$ is the unit vector pointing from the proton to the electron. Here $\mathbf{K}_{\rm e}$ acts on the electron part and $\mathbf{S}_{\rm e}$ on the proton part. Both $\mathbf{K}_{\rm e}$ and $\mathbf{S}_{\rm e}$ are tensor operators of rank one. Using first order perturbation theory, we want to show that matrix elements of this Hamiltonian in the coupled states:

$$|\,n,(\ell,s_e)\,j,s_p,f,m_f\,\rangle\,, \tag{21.327}$$

are diagonal for states with the same value of j, and we want to find the splitting energy. We first want to write $\mathbf{S}_{\rm e}-3\,(\mathbf{S}_{\rm e}\cdot\hat{\mathbf{R}})\,\hat{\mathbf{R}}$ as a tensor operator. We state the result of this derivation as the following theorem:

Theorem 52. *The vector* $\mathbf{S}_e-3\,(\mathbf{S}_e\cdot\hat{\mathbf{R}})\,\hat{\mathbf{R}}$ *can be written as a rank one tensor operator of the form:*

$$[\,\mathbf{S}_e-3\,(\mathbf{S}_e\cdot\hat{\mathbf{R}})\,\hat{\mathbf{R}}\,]_{1,q} = \sqrt{10}\,[\,C_2(\hat{\mathbf{R}})\otimes S_1(\mathbf{e})\,]_{1,q}\,. \tag{21.328}$$

Proof. We start by writing:

$$[\,\mathbf{S}_{\rm e}-3\,(\mathbf{S}_{\rm e}\cdot\hat{\mathbf{R}})\,\hat{\mathbf{R}}\,]_{1,q} = \sum_{q_1} S_{1,q_1}(\mathbf{e})\,\{\,\delta_{q_1,q}-3\,C^*_{1,q_1}(\hat{\mathbf{R}})\,C_{1,q}(\hat{\mathbf{R}})\,\}\,. \tag{21.329}$$

Next, we have:

$$\begin{aligned} 3\,C^*_{1,q_1}(\hat{\mathbf{R}})\,C_{1,q}(\hat{\mathbf{R}}) &= 3\,(-)^{q_1}\sum_{k,q_2} C_{k,q_2}(\hat{\mathbf{R}})\,\langle\,1,-q_1,1,q\,|\,(1,1)\,k,q\,\rangle\,\langle\,1,0,1,0\,|\,(1,1)\,k,0\,\rangle \\ &= \delta_{q_1,q} - \sqrt{10}\sum_{q_2} C_{2,q_2}(\hat{\mathbf{R}})\,\langle\,2,q_2,1,q_1\,|\,(1,2)\,1,q\,\rangle\,. \end{aligned} \tag{21.330}$$

Here we have used $\langle\,1,0,1,0\,|\,(1,1)\,k,0\,\rangle = -1/\sqrt{3}$, 0, and $+\sqrt{2/3}$ for $k = 0$, 1, and 2 respectively. Substitution of (21.330) into (21.329) gives:

$$[\,\mathbf{S}_{\rm e} - 3\,(\mathbf{S}_{\rm e}\cdot\hat{\mathbf{R}})\,\hat{\mathbf{R}}\,]_{1,q} = \sqrt{10}\sum_{q_2}\langle\,2,q_2,1,q_1\,|\,(1,2)\,1,q\,\rangle\,C_{2,q_2}(\hat{\mathbf{R}})\,S_{1,q_1}(\mathbf{e}) = \sqrt{10}\,[\,C_2(\hat{\mathbf{R}})\otimes S_1(\mathbf{e})\,]_{1,q}, \tag{21.331}$$

which proves the theorem. □

We now want to find the matrix elements of the scalar product:

$$\langle\,n,(\ell,s_e)\,j,s_p,f,m_f\,|\,[K_1(\mathrm{e})\odot S_1(\mathrm{p})\,]\,|\,n,(\ell',s_e)\,j,s_p,f',m'_f\,\rangle\,, \tag{21.332}$$

where $K_{1,q}(\mathrm{e})$ is the rank one tensor operator:

$$K_{1,q}(\mathrm{e}) = L_{1,q}(\mathrm{e}) - \sqrt{10}\,[\,C_2(\hat{\mathbf{R}})\otimes S_1(\mathbf{e})\,]_{1,q}\,. \tag{21.333}$$

Here $K_1(\mathrm{e})$ only operates on the electron part (the first part of the coupled state) and $S_1(\mathrm{p})$ on the proton part (the second part of the coupled state). So using Theorem 48, we find:

$$\begin{aligned} &\langle\,n,(\ell,s_e)\,j,s_p,f,m_f\,|\,[K_1(\mathrm{e})\odot S_1(\mathrm{p})\,]\,|\,n,(\ell',s_e)\,j,s_p,f',m'_f\,\rangle/\hbar^2 \\ &\quad= \delta_{f,f'}\,\delta_{m_f,m'_f}\,(-)^{j+s_p+f}\begin{Bmatrix} f & s_p & j \\ 1 & j & s_p \end{Bmatrix}\langle\,(\ell,s_e)\,j\,\|\,K_1(\mathrm{e})\,\|\,(\ell',s_e)\,j\,\rangle\,\langle\,s_p\,\|\,S_1(\mathrm{p})\,\|\,s_p\,\rangle/\hbar^2 \\ &\quad= \delta_{f,f'}\,\delta_{m_f,m'_f}\,(-)^{f+j+1/2}\,\sqrt{3/2}\begin{Bmatrix} f & 1/2 & j \\ 1 & j & 1/2 \end{Bmatrix}\langle\,(\ell,s_e)\,j\,\|\,K_1(\mathrm{e})\,\|\,(\ell',s_e)\,j\,\rangle/\hbar \\ &\quad= \delta_{f,f'}\,\delta_{m_f,m'_f}\,\frac{f(f+1)-j(j+1)-3/4}{2\sqrt{j(j+1)(2j+1)}}\,\langle\,(\ell,s_e)\,j\,\|\,K_1(\mathrm{e})\,\|\,(\ell',s_e)\,j\,\rangle/\hbar\,. \end{aligned} \tag{21.334}$$

Since $L_1(\mathrm{e})$ only operates on the first part of the coupled scheme $(\ell,s_e)\,j$, its reduced matrix elements can be found by application of Theorem 49, and we find:

$$\begin{aligned} \langle\,(\ell,s_e)\,j\,\|\,L_1(\mathrm{e})\,\|\,(\ell',s_e)\,j\,\rangle/\hbar &= (-)^{\ell+j+3/2}\,(2j+1)\begin{Bmatrix} \ell & j & 1/2 \\ j & \ell' & 1 \end{Bmatrix}\langle\,\ell\,\|\,L_1(\mathrm{e})\,\|\,\ell'\,\rangle/\hbar\,. \\ &= \delta_{\ell,\ell'}\,\frac{1}{2}\,\sqrt{2\ell(2\ell+1)(2\ell+2)}\,(-)^{\ell+j+3/2}\,(2j+1)\begin{Bmatrix} 1/2 & j & \ell \\ 1 & \ell & j \end{Bmatrix} \\ &= \delta_{\ell,\ell'}\,\frac{1}{2}\,\sqrt{\frac{2j+1}{j(j+1)}}\,\{\,j(j+1)+\ell(\ell+1)-3/4\,\}\,, \end{aligned} \tag{21.335}$$

where we have used Table 21.3. Using Theorem 47 the reduced matrix element of $\sqrt{10}\,[\,C_2(\hat{\mathbf{r}})\otimes S_1(\mathbf{e})\,]_1$ is given by

$$\begin{aligned} &\sqrt{10}\,\langle\,(\ell,s_e)\,j\,\|\,[\,C_2(\hat{\mathbf{R}})\otimes S_1(\mathbf{e})\,]_1\,\|\,(\ell',s_e)\,j\,\rangle/\hbar \\ &\quad= \sqrt{30}\,(2j+1)\begin{Bmatrix} \ell & \ell' & 2 \\ s_e & s_e & 1 \\ j & j & 1 \end{Bmatrix}\langle\,\ell\,\|\,C_2(\hat{\mathbf{r}})\,\|\,\ell'\,\rangle\,\langle\,s_e\,\|\,S_1(\mathbf{e})\,\|\,s_e\,\rangle/\hbar \\ &\quad= (-)^{\ell}\,3\sqrt{5}\,(2j+1)\,\sqrt{(2\ell+1)(2\ell'+1)}\begin{pmatrix} \ell & 2 & \ell' \\ 0 & 0 & 0 \end{pmatrix}\begin{Bmatrix} \ell & \ell' & 2 \\ 1/2 & 1/2 & 1 \\ j & j & 1 \end{Bmatrix}\,. \end{aligned} \tag{21.336}$$

The $6j$-symbol vanished unless $\ell + \ell' + 2$ is even. But since we are only considering states with the same j values, this means that $\ell = \ell'$. From tables in Edmonds, we have:

$$\begin{pmatrix} \ell & 2 & \ell \\ 0 & 0 & 0 \end{pmatrix} = (-)^{\ell+1} \sqrt{\frac{\ell(\ell+1)}{(2\ell-1)(2\ell+1)(2\ell+3)}}\,, \tag{21.337}$$

and from tables in Matsunobu and Takebe [18], we have:

$$\begin{Bmatrix} \ell & \ell & 2 \\ j & j & 1 \\ 1/2 & 1/2 & 1 \end{Bmatrix} = \frac{1}{3\sqrt{5}\,(2j+1)(2\ell+1)} \begin{cases} (-)\,\sqrt{2\ell(2\ell-1)}\,, & \text{for } j = \ell + 1/2, \\ (+)\,\sqrt{(2\ell+1)(2\ell+3)}\,, & \text{for } j = \ell - 1/2. \end{cases} \tag{21.338}$$

Putting Eqs. (21.337) and (21.338) into Eq. (21.336) gives:

$$\begin{aligned} \sqrt{10}\,\langle\,(\ell,s_e)\,j\,\|\,[\,C_2(\hat{\mathbf{R}}) \otimes S_1(\mathbf{e})\,]_1\,\|\,(\ell',s_e)\,j\,\rangle/\hbar & \\ = \delta_{\ell,\ell'}\,\frac{1}{2}\,\sqrt{\frac{2j+1}{j(j+1)}} \times \begin{cases} \ell & \text{for } j = \ell + 1/2, \\ -(\ell+1) & \text{for } j = \ell - 1/2 \end{cases} & \\ = \delta_{\ell,\ell'}\,\frac{1}{2}\,\sqrt{\frac{2j+1}{j(j+1)}}\,\{\,j(j+1) - \ell(\ell+1) - 3/4\,\}\,. & \end{aligned} \tag{21.339}$$

So subtracting (21.339) from (21.335), we find the result:

$$\langle\,(\ell,s_e)\,j\,\|\,K_1(\mathrm{e})\,\|\,(\ell',s_e)\,j\,\rangle/\hbar = \delta_{\ell,\ell'}\,\ell(\ell+1)\,\sqrt{\frac{2j+1}{j(j+1)}}\,. \tag{21.340}$$

Putting this into Eq. (21.334) gives:

$$\begin{aligned} \langle\,n,(\ell,s_e)\,j,s_p,f,m_f\,|\,[K_1(\mathrm{e}) \odot S_1(\mathrm{p})\,]\,|\,n,(\ell',s_e)\,j,s_p,f',m'_f\,\rangle/\hbar^2 & \\ = \delta_{f,f'}\,\delta_{m_f,m'_f}\,\delta_{\ell,\ell'}\,\frac{\ell(\ell+1)}{2j(j+1)}\,\{\,f(f+1) - j(j+1) - 3/4\,\}\,. & \end{aligned} \tag{21.341}$$

So we have shown that:

$$\langle\,n,(\ell,s_e)\,j,s_p,f,m_f\,|\,\bar{H}_{\mathrm{hf}}\,|\,n,(\ell',s_e)\,j,s_p,f',m'_f\,\rangle = \delta_{f,f'}\,\delta_{m_f,m'_f}\,\delta_{\ell,\ell'}\,\Delta E_{n,\ell,j,f}\,, \tag{21.342}$$

where, in atomic units, the energy shift $\Delta\bar{E}_{n,\ell,j,f}$ is given by:

$$\Delta\bar{E}_{n,\ell,j,f} = 2\lambda_{\mathrm{p}}\left(\frac{m}{M}\right)\alpha^2\,\frac{f(f+1) - j(j+1) - 3/4}{n^3\,j(j+1)\,(2\ell+1)}\,. \tag{21.343}$$

Here we have used:

$$\left\langle \frac{1}{\bar{R}^3} \right\rangle_{n,\ell} = \frac{2}{n^3\,\ell(\ell+1)(2\ell+1)}\,. \tag{21.344}$$

Eq. (21.343) is quoted in our discussion of the hyperfine structure of hydrogen in Section 22.3.8.

21.6.4 The Zeeman effect in hydrogen

The Hamiltonian for the Zeeman effect in Hydrogen is given by Eq. (22.191), where we found:

$$H_{\mathrm{z}} = \mu_B\,(\,\mathbf{L} + 2\,\mathbf{S}\,) \cdot \mathbf{B}/\hbar\,, \qquad \text{with} \qquad \mu_B = \frac{e\,\hbar}{2\,mc}\,. \tag{21.345}$$

We shall find matrix elements within the hyperfine splitting levels. That is, taking the z-axis in the direction of the $\mathbf{B}$ field,

$$\langle\,(\ell,s_e)\,j,s_p,f,m_f\,|\,H_z\,|\,(\ell,s_e)\,j,s_p,f',m'_f\,\rangle$$
$$=\mu_B\,B\,\langle\,(\ell,s_e)\,j,s_p,f,m_f\,|\,(\,L_z+2\,S_z\,)\,|\,(\ell,s_e)\,j,s_p,f',m'_f\,\rangle/\hbar\,. \quad (21.346)$$

Now both L_z and S_z are $q=0$ components of tensor operators of rank $k=1$. So using the Wigner-Eckart Theorem 44, and Theorems 49 and 50, we find:

$$\langle\,(\ell,s_e)\,j,s_p,f,m_f\,|\,L_{1,0}(e)\,|\,(\ell,s_e)\,j,s_p,f',m'_f\,\rangle/\hbar$$
$$=(-)^{f-m_f}\begin{pmatrix}f&1&f'\\-m_f&0&m'_f\end{pmatrix}\langle\,(\ell,s_e)\,j,s_p,f\,\|\,L_1(e)\,\|\,(\ell,s_e)\,j,s_p,f'\,\rangle/\hbar$$
$$=(-)^{f-m_f+j+1/2+f'+1}\sqrt{(2f+1)(2f'+1)}\begin{pmatrix}f&1&f'\\-m_f&0&m'_f\end{pmatrix}\begin{Bmatrix}j&f&1/2\\f'&j&1\end{Bmatrix}\langle\,(\ell,s_e)\,j\,\|\,L_1(e)\,\|\,(\ell,s_e)\,j\,\rangle/\hbar$$
$$=(-)^{f-m_f+j+1/2+f'+j+\ell+1/2}\,(2j+1)\sqrt{(2f+1)(2f'+1)}\begin{pmatrix}f&1&f'\\-m_f&0&m'_f\end{pmatrix}\begin{Bmatrix}j&f&1/2\\f'&j&1\end{Bmatrix}$$
$$\times\begin{Bmatrix}\ell&j&1/2\\j&\ell&1\end{Bmatrix}\langle\,\ell\,\|\,L_1(e)\,\|\,\ell\,\rangle/\hbar$$
$$=(-)^{f-m_f+j+1/2+f'+j+\ell+1/2}\,(2j+1)\sqrt{(2f+1)(2f'+1)\,\ell(\ell+1)(2\ell+1)}\begin{pmatrix}f&1&f'\\-m_f&0&m'_f\end{pmatrix}\begin{Bmatrix}j&f&1/2\\f'&j&1\end{Bmatrix}$$
$$\times\begin{Bmatrix}\ell&j&1/2\\j&\ell&1\end{Bmatrix}$$
$$=(-)^{f+f'-m_f+j-1/2}\,\frac{1}{2}\sqrt{\frac{(2j+1)(2f+1)(2f'+1)}{j(j+1)}}\begin{pmatrix}f&1&f'\\-m_f&0&m'_f\end{pmatrix}\begin{Bmatrix}j&f&1/2\\f'&j&1\end{Bmatrix}$$
$$\times\left[\,j(j+1)+\ell(\ell+1)-3/4\,\right] \quad (21.347)$$

and

$$\langle\,(\ell,s_e)\,j,s_p,f,m_f\,|\,S_{1,0}(e)\,|\,(\ell,s_e)\,j,s_p,f',m'_f\,\rangle/\hbar$$
$$=(-)^{f-m_f}\begin{pmatrix}f&1&f'\\-m_f&0&m'_f\end{pmatrix}\langle\,(\ell,s_e)\,j,s_p,f\,\|\,S_1(e)\,\|\,(\ell,s_e)\,j,s_p,f'\,\rangle/\hbar$$
$$=(-)^{f-m_f+j+1/2+f'+1}\sqrt{(2f+1)(2f'+1)}\begin{pmatrix}f&1&f'\\-m_f&0&m'_f\end{pmatrix}\begin{Bmatrix}j&f&1/2\\f'&j&1\end{Bmatrix}\langle\,(\ell,s_e)\,j\,\|\,S_1(e)\,\|\,(\ell,s_e)\,j\,\rangle/\hbar$$
$$=(-)^{f-m_f+j+1/2+f'+j+\ell+1/2}\,(2j+1)\sqrt{(2f+1)(2f'+1)}\begin{pmatrix}f&1&f'\\-m_f&0&m'_f\end{pmatrix}\begin{Bmatrix}j&f&1/2\\f'&j&1\end{Bmatrix}$$
$$\times\begin{Bmatrix}1/2&j&\ell\\j&1/2&1\end{Bmatrix}\langle\,1/2\,\|\,S_1(e)\,\|\,1/2\,\rangle/\hbar$$

$$=(-)^{f-m_f+j+1/2+f'+j+\ell+1/2}\,(2j+1)\sqrt{(2f+1)(2f'+1)\,3/2}\begin{pmatrix}f&1&f'\\-m_f&0&m'_f\end{pmatrix}\begin{Bmatrix}j&f&1/2\\f'&j&1\end{Bmatrix}$$
$$\times\begin{Bmatrix}1/2&j&\ell\\j&1/2&1\end{Bmatrix}=(-)^{f+f'-m_f+j-1/2}\,\frac{1}{2}\sqrt{\frac{(2j+1)(2f+1)(2f'+1)}{j(j+1)}}\begin{pmatrix}f&1&f'\\-m_f&0&m'_f\end{pmatrix}\begin{Bmatrix}j&f&1/2\\f'&j&1\end{Bmatrix}$$
$$\times\left[\,j(j+1)-\ell(\ell+1)+3/4\,\right]. \quad (21.348)$$

So multiplying Eq. (21.348) by a factor of two and adding it to Eq. (21.347) gives:

$$\begin{aligned}\langle\,(\ell,s_e)\,j,s_p,f,m_f\,|\,H_z\,|\,(\ell,s_e)\,j,s_p,f',m'_f\,\rangle\\ =\mu_B\,B\,(-)^{f+f'-m_f+j-1/2}\,\frac{1}{2}\sqrt{\frac{(2j+1)(2f+1)(2f'+1)}{j(j+1)}}\begin{pmatrix}f&1&f'\\-m_f&0&m'_f\end{pmatrix}\begin{Bmatrix}j&f&1/2\\f'&j&1\end{Bmatrix}\\ \times\left[3j(j+1)-\ell(\ell+1)+3/4\right].\end{aligned}\tag{21.349}$$

The $3j$-symbol vanishes unless $m'_f=m_f$, so the matrix element connects only states of the same m_f. Now if $f'=f$, we find the simple result:

$$\begin{aligned}\langle\,(\ell,s_e)\,j,s_p,f,m_f\,|\,H_z\,|\,(\ell,s_e)\,j,s_p,f,m_f\,\rangle\\ =(\,\mu_B\,B\,)\,m_f\,\frac{\left[f(f+1)+j(j+1)-3/4\right]\left[3j(j+1)-\ell(\ell+1)+3/4\right]}{4\,f(f+1)\,j(j+1)}.\end{aligned}\tag{21.350}$$

On the other hand, if $f'=f+1$, we get:

$$\begin{aligned}\langle\,(\ell,s_e)\,j,s_p,f,m_f\,|\,H_z\,|\,(\ell,s_e)\,j,s_p,f+1,m_f\,\rangle\\ =(\,\mu_B\,B\,)\,\frac{3j(j+1)-\ell(\ell+1)+3/4}{j(j+1)\,(f+1)}\\ \times\sqrt{\frac{(f-m_f+1)(f+m_f+1)(f+j+5/2)(f+j+1/2)(f-j+3/2)(j-f+1/2)}{(2f+1)(2f+3)}}.\end{aligned}\tag{21.351}$$

with an identical expression for the matrix elements of $\langle\,(\ell,s_e)\,j,s_p,f+1,m_f\,|\,H_z\,|\,(\ell,s_e)\,j,s_p,f,m_f\,\rangle$. We use these results in Section 22.3.9.

21.6.5 The Stark effect in hydrogen

In Section 22.3.10 we derived a Hamiltonian for the Stark effect in hydrogen. We found:

$$H_S=e\,a\,E_0\,\bar{R}\,C_{1,0}(\Omega)\,.\tag{21.352}$$

So we need to find the matrix elements:

$$\begin{aligned}\langle\,(\ell,s)\,j,m_j\,|\,C_{1,0}(\Omega)\,|\,(\ell',s')\,j',m'_j\,\rangle=(-)^{j-m_j}\begin{pmatrix}j&1&j'\\-m_j&0&m'_j\end{pmatrix}\langle\,(\ell,s)\,j\,\|\,C_1(\Omega)\,\|\,(\ell',s')\,j'\,\rangle\\ =\delta_{s,s'}\,\delta_{m_j,m'_j}\,(-)^{j-m_j+\ell+1/2+j'+1}\sqrt{(2j+1)(2j'+1)}\begin{pmatrix}j&1&j'\\-m_j&0&m_j\end{pmatrix}\begin{Bmatrix}\ell&j&1/2\\j'&\ell'&1\end{Bmatrix}\langle\,\ell\,\|\,C_1(\Omega)\,\|\,\ell'\,\rangle\\ =\delta_{s,s'}\,\delta_{m_j,m'_j}\,(-)^{j+j'-m_j-1/2}\sqrt{(2j+1)(2j'+1)(2\ell+1)(2\ell'+1)}\begin{pmatrix}j&1&j'\\-m_j&0&m_j\end{pmatrix}\begin{pmatrix}\ell&1&\ell'\\0&0&0\end{pmatrix}\\ \times\begin{Bmatrix}\ell&j&1/2\\j'&\ell'&1\end{Bmatrix}.\end{aligned}\tag{21.353}$$

Now since $\ell+\ell'$ must be odd, the diagonal matrix elements all vanish. For our case $j=\ell\pm1/2$ and $j'=\ell'\pm1/2$, so the only contributing non-zero elements are those in which $\ell'=\ell\pm1$ *and* $j'=j\pm1$ Rather than find a general formula, it is simpler to just work out the matrix elements for cases we want. For the $n=1$ fine structure levels, there is only the $1s_{1/2}$ state, so the matrix element vanishes for this case. For the $n=2$ fine structure levels, we have three states: $2s_{1/2}$, $2p_{1/2}$, and $2p_{3/2}$. For these cases, we find:

$$\begin{aligned}\langle\,2s_{1/2},m\,|\,C_{1,0}(\Omega)\,|\,2p_{1/2},m\,\rangle&=-\frac{2}{3}\,m\,,\\ \langle\,2s_{1/2},m\,|\,C_{1,0}(\Omega)\,|\,2p_{3/2},m\,\rangle&=\frac{1}{3}\sqrt{(3/2-m)(3/2+m)}=\frac{\sqrt{2}}{3}\,,\end{aligned}\tag{21.354}$$

for $m = \pm 1/2$. We will use these results in Section 22.3.10.

21.6.6 Matrix elements of two-body nucleon-nucleon potentials

In the nuclear shell model, nucleons (protons and neutrons) with spin $s = 1/2$ are in $(\ell, s)\, j, m_j$ coupled orbitals with quantum numbers given by: $n(\ell)_j = 1s_{1/2}, 1p_{1/2}, 2s_{1/2}, 2p_{3/2}, \cdots$. We leave it to a nuclear physics book to explain why this is often a good approximation (see, for example, the book *Nuclear Physics* by J. D. Walecka). The nucleon-nucleon interaction between nucleons in these orbitals give a splitting of the shell energies of the nucleus. One such interaction is the one-pion exchange interaction of the form:

$$V(\mathbf{r}_1, \mathbf{r}_2) = V_0 \frac{e^{-\mu r}}{r} \left\{ \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 + \left[\frac{1}{(\mu r)^2} + \frac{1}{(\mu r)} + \frac{1}{3} \right] S_{1,2} \right\} \boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2 \,, \tag{21.355}$$

where $r = |\mathbf{r}_1 - \mathbf{r}_2|$ is the distance between the nucleons, $\mu = m_\pi c/\hbar$ the inverse pion Compton wavelength, $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ the spin operators, $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ the isospin operators for the two nucleons, and $S_{1,2}$ the tensor operator:

$$S_{1,2} = 3\,(\hat{\mathbf{r}} \cdot \boldsymbol{\sigma}_1)\,(\hat{\mathbf{r}} \cdot \boldsymbol{\sigma}_2) - \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 \,. \tag{21.356}$$

The nuclear state is given by the coupling:

$$|\, n_1, n_2; (\ell_1, s_1)\, j_1, (\ell_2, s_2)\, j_2, j, m \,\rangle \tag{21.357}$$

To find the nuclear energy levels, we will need to find matrix elements of the nuclear force between these states. The calculation of these matrix elements generally involve a great deal of angular momentum technology. The nucleon-nucleon force, given in Eq. (21.355), is only one example of a static nucleon-nucleon interaction. Other examples are the V_6 and V_{12} Argonne interactions. Matrix elements of these interactions have been worked out in the literature by B. Mihaila and J. Heisenberg [21]. We show how to compute some of these matrix elements here.

Scalar force

Let us first make a multipole expansion of a scalar potential. Let $\mathbf{r}_1$ and $\mathbf{r}_2$ be the location of nucleon 1 and nucleon 2 in the center of mass coordinate system of the nucleus. Then a scalar potential, which depends only on the magnitude of the distance between the particles is given by:

$$\begin{aligned} V_S(\mathbf{r}_1, \mathbf{r}_2) &= V_S(|\mathbf{r}_1 - \mathbf{r}_2|) = V(r_1, r_2, \cos\theta) \\ &= \sum_{k=0}^{\infty} V_k(r_1, r_2)\, P_k(\cos\theta) = \sum_{k=0}^{\infty} V_k(r_1, r_2)\, [\, C_k(\Omega_1) \odot C_k(\Omega_2)\,] \,, \end{aligned} \tag{21.358}$$

where

$$V_k(r_1, r_2) = \frac{2k+1}{2} \int_{-1}^{+1} V(r_1, r_2, \cos\theta)\, P_k(\cos\theta)\, \mathrm{d}(\cos\theta) \,, \tag{21.359}$$

and where we have used Eq. (21.279). Eq. (21.358) is now in the form required for the j-j coupling state given in (21.357). So now applying Theorem 21.310, we find:

$$\begin{aligned} \Delta E &= \langle\, n_1, n_2; (\ell_1, s_1)\, j_1, (\ell_2, s_2)\, j_2, j, m \,|\, V(|\mathbf{r}_1 - \mathbf{r}_2|) \,|\, n_1, n_2; (\ell_1, s_1)\, j_1, (\ell_2, s_2)\, j_2, j', m' \,\rangle \\ &= \sum_{k=0}^{\infty} F_k(1,2)\, \langle\, (\ell_1, s_1)\, j_1, (\ell_2, s_2)\, j_2, j, m \,|\, [\, C_k(\Omega_1) \odot C_k(\Omega_2)\,] \,|\, (\ell_1, s_1)\, j_1, (\ell_2, s_2)\, j_2, j, m \,\rangle \\ &= \delta_{j,j'}\, \delta_{m,m'} \sum_{k=0}^{\infty} F_k(1,2)\, (-)^{j_1+j_2+j} \begin{Bmatrix} j & j_2 & j_1 \\ k & j_1 & j_2 \end{Bmatrix} \\ &\qquad \times \langle\, (\ell_1, s_1)\, j_1 \,\|\, C_k(\Omega_1) \,\|\, (\ell_1, s_1)\, j_1 \,\rangle\, \langle\, (\ell_2, s_2)\, j_2 \,\|\, C_k(\Omega_2) \,\|\, (\ell_2, s_2)\, j_2 \,\rangle \,. \end{aligned} \tag{21.360}$$

Here

$$F_k(1,2) = \int_0^\infty r_1^2\,\mathrm{d}r_1 \int_0^\infty r_2^2\,\mathrm{d}r_2\, R^2_{n_1,\ell_1,j_1}(r_1)\, R^2_{n_2,\ell_2,j_2}(r_2)\, V_k(r_1,r_2) \tag{21.361}$$

are integrals over the radial wave functions for the nucleons in the orbitals $n_1(\ell_1)_{j_1}$ and $n_2(\ell_2)_{j_2}$. It is now a simple matter to compute the reduced matrix elements of $C_k(\Omega)$ using Theorem 49 and Eqs. (21.220) and (21.286). We find:

$$\begin{aligned}\langle\,(\ell,1/2)\,j\,\|\,C_k(\Omega)\,\|\,(\ell',1/2)\,j'\,\rangle &= (-)^{\ell+\ell'+j'+k}\sqrt{(2j+1)(2j'+1)}\begin{Bmatrix}\ell & j & 1/2\\ j' & \ell' & k\end{Bmatrix}\langle\,\ell\,\|\,C_k\,\|\,\ell'\,\rangle\\ &= (-)^{\ell'+j'+k}\sqrt{(2j+1)(2j'+1)(2\ell+1)(2\ell'+1)}\begin{Bmatrix}\ell & j & 1/2\\ j' & \ell' & k\end{Bmatrix}\begin{pmatrix}\ell & k & \ell'\\ 0 & 0 & 0\end{pmatrix}\\ &= (-)^{k+j'-1/2}\sqrt{(2j+1)(2j'+1)}\begin{Bmatrix}j & j' & k\\ 1/2 & -1/2 & 0\end{Bmatrix}\delta(\ell,\ell',k)\,,\end{aligned} \tag{21.362}$$

where $\delta(\ell,\ell',k) = 1$ if $\ell+\ell'+k$ is even and (ℓ,ℓ',k) satisfy the triangle inequality, otherwise it is zero. Substitution into Eq. (21.360) gives $\Delta E = \delta_{j,j'}\delta_{m,m'}\,E_j$, where E_j is given by:

$$\begin{aligned}\Delta E_j = \sum_{k=0}^{\infty} F_k(1,2)\,(-)^{j+1}\,(2j_1+1)(2j_2+1)\\ \times\begin{Bmatrix}j & j_2 & j_1\\ k & j_1 & j_2\end{Bmatrix}\begin{Bmatrix}j_1 & j_1 & k\\ 1/2 & -1/2 & 0\end{Bmatrix}\begin{Bmatrix}j_2 & j_2 & k\\ 1/2 & -1/2 & 0\end{Bmatrix}\delta(\ell_1,\ell_1,k)\,\delta(\ell_2,\ell_2,k)\,,\end{aligned} \tag{21.363}$$

which completes the calculation. Note that k has to be even.

Exercise 71. If $j_1 = j_2$ and all values of $F_k(1,2)$ are negative corresponding to an attractive nucleon-nucleon potential, show that the expected nuclear spectra is like that shown in Fig. **??**. [J. D. Walecka, p. 517].

Spin-exchange force

The nucleon-nucleon spin-exchange force is of the form:

$$V_{SE}(\mathbf{r}_1,\mathbf{r}_2,\sigma_1,\sigma_2) = V_{SE}(|\mathbf{r}_1-\mathbf{r}_2|)\,\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2 = \sum_{k,\ell}(-)^{\ell+1-k}\,V_\ell(r_1,r_2)\,[\,T_{\ell,k}(1)\odot T_{\ell,k}(2)\,]\,, \tag{21.364}$$

where

$$\begin{aligned}T_{(\ell,1)\,k,q}(1) &= [\,C_\ell(\Omega_1)\otimes\sigma(1)\,]_{k,q}\,,\\ T_{(\ell,1)\,k,q}(2) &= [\,C_\ell(\Omega_2)\otimes\sigma(2)\,]_{k,q}\,.\end{aligned} \tag{21.365}$$

This now is in a form suitable for calculation in j-j coupling.

Spin-orbit force

XXX

Tensor force

The tensor force is of the form:

$$\begin{aligned}V_T(\mathbf{r}_1,\mathbf{r}_2,\sigma_1,\sigma_2) &= V_T(|\mathbf{r}_1-\mathbf{r}_2|)\,\{\,(\,\boldsymbol{\sigma}_1\cdot\hat{\mathbf{r}}_{12}\,)\,(\,\boldsymbol{\sigma}_2\cdot\hat{\mathbf{r}}_{12}\,) - (\,\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2\,)/3\,\}\\ &= V_T(|\mathbf{r}_1-\mathbf{r}_2|)\,[\,L_2(1,2)\odot S_2(1,2)]\,,\end{aligned} \tag{21.366}$$

where

$$\begin{aligned} S_{2,q}(1,2) &= [\,\sigma_1(1)\otimes\sigma_1(2)\,]_{2,q}\,, \\ L_{2,q}(1,2) &= [\,\hat{R}_1(1,2)\otimes\hat{R}_1(1,2)\,]_{2,q}\,, \end{aligned} \tag{21.367}$$

with $\hat{R}_1(1,2)$ the spherical vector of components of the unit vector $\mathbf{r}_{12}$. We follow the method described by de-Shalit and Walecka [22] here. Expanding

$$V_T(|\mathbf{r}_1-\mathbf{r}_2|) = \sum_{k=0}^{\infty} V_{T\,k}(r_1,r_2)\,[\,C_k(\Omega_1)\odot C_k(\Omega_2)\,]\,, \tag{21.368}$$

After some work, we find:

$$V_T(\mathbf{r}_1,\mathbf{r}_2,\sigma_1,\sigma_2) = \sum_{k,\ell}(-)^{\ell+1-k}\,V_\ell(r_1,r_2)\,[\,X_{\ell,k}(1)\odot X_{\ell,k}(2)\,]\,, \tag{21.369}$$

where

21.6.7 Density matrix for the Deuteron

References

[1] L. C. Biedenharn and H. van Dam, *Quantum theory of angular momentum*, Perspectives in physics (Academic Press, New York, NY, 1965).

ANNOTATION: This is a collection of early papers on the theory of angular momentum.

[2] A. R. Edmonds, *Angular momentum in quantum mechanics* (Princeton University Press, Princeton, NJ, 1996), fourth printing with corrections, second edition.

ANNOTATION: This printing corrected several major errors in Chapter 4 in earlier printings.

[3] E. U. Condon and G. H. Shortley, *The theory of atomic spectra* (Cambridge University Press, New York, 1953).

[4] G. Racah, "Theory of complex spectra II," Phys. Rev. **62**, 438 (1942).

[5] L. C. Biedenharn and J. D. Louck, *Angular momentum in quantum physics: theory and application*, volume 8 of *Encyclopedia of mathematics and its applications* (Addison-Wesley, Reading, MA, 1981).

[6] M. A. Rose, *Elementary theory of angular momentum* (John Wiley & Sons, New York, NY, 1957).

[7] E. Merzbacher, *Quantum Mechanics* (John Wiley & Sons, New York, NY, 1970), second edition.

[8] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics* (Addison-Wesley, Reading, MA, 2002), third edition.

[9] W. R. Hamilton, *Lectures on Quaternions* (Dublin, 1853).

[10] F. Klein, "Über binäre Formen mit linearen Transformationsen in sich selbst," Math. Ann. **9**, 183 (1875).

[11] A. Klein, *The Icosahedron* (Dover, 1956).

ANNOTATION: Reproduction of Klein's original book published in 1884.

[12] A. Cayley, "On the correspondance of homographies and rotations," Math. Ann. **15**, 238 (1879).

ANNOTATION: reprinted in Collected Math Papers X, pp. 153–154.

[13] A. Cayley, "On the application of quaternions to the theory of rotations," Phil. Mag. **3**, 196 (1848).

ANNOTATION: reprinted in Collected Math. Papers I, pp. 405–409.

[14] M. Bouten, "On the rotation operators in quantum mechanics," Physica **42**, 572 (1969).

[15] A. A. Wolf, "Rotation operators," Am. J. Phys. **37**, 531 (1969).

[16] M. Rotenberg, R. Bivins, N. Metropolis, and J. J. K. Wooten, *The 3-j and 6-j symbols* (The Technology Press, MIT, Cambridge, MA, 1959).

ANNOTATION: The introduction to these tables gives many relations for Clebch-Gordan coefficients, 3-j and 6-j symbols. The tables are in powers of prime notation.

[17] D. M. Brink and G. R. Satchler, *Angular momentum* (Clarendon Press, Oxford, England, 1968), second edition.

ANNOTATION: This book contanes an excellent appendix containing many realtions between $3j$, $6j$, and $9j$ symbols. Brink and Satchler define a reduced matrix element which is related to ours by: $\langle\, j \,\|\, T \,\|\, j' \,\rangle = \sqrt{2j+1}\,[\,\langle\, j \,\|\, T \,\|\, j' \,\rangle\,]_{\text{Brink-Satchler}}$. Also they use an active rotation matrix, which differs from our convention here.

[18] H. Matsunobu and H. Takebe, "Tables of U coeficients," Prog. Theo. Physics **14**, 589 (1955).

[19] E. P. Wigner, *Gruppentheorie und ihre Anwendung auf dei Quantenmechanic der Atomspektren* (Braunschweig, Berlin, 1931). English translation: Academic Press, Inc, New York, 1959.

[20] C. Eckart, "The application of group theory to the quantum dynamics of monatomic systems," Revs. Mod. Phys. **2**, 305 (1930).

[21] B. Mihaila and J. Heisenberg, "Matrix elements of the Argonne V_6 and V_{12} potentials," .

[22] A. de Shalit and J. D. Walecka, "Spectra of odd nuclei," Nuc. Phys. **22**, 184 (1961).

Chapter 22

Non-relativistic electrodynamics

In this chapter, we first derive the Hamiltonian and equations of motion for a non-relativistic charged particle interacting with an external electromagnetic field. After discussing the motion of a charged particle in a constant electric field, we find the eigenvalues and eigenvectors for the hydrogen atom. We then discuss the fine structure, the hyperfine structure, the Zeeman and Stark effects in hydrogen.

22.1 The Lagrangian

We start with the classical Lagrangian of a non-relativistic particle of mass m and charge q interacting with an external electromagnetic field. The justification of this Lagrangian is that it reproduces Newton's laws with a Lorentz force for the electromagnetic field, as we shall see. This Lagrangian is given by:

$$L(\mathbf{r},\mathbf{v}) = \frac{1}{2}\,m\,\mathbf{v}^2 - q\,\phi(\mathbf{r},t) + q\,\mathbf{v}\cdot\mathbf{A}(\mathbf{r},t)/c\,, \tag{22.1}$$

where $\mathbf{v} = \dot{\mathbf{r}}$ and $\phi(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$ are the external electromagnetic potential fields which transform under rotations as scalar and vectors. We use electrostatic units in this section so that, for example, the charge of the electron is $q = -e = -4.511\times10^{-10}$ esu. The appearance of the velocity of light here is because of the units we use; however, there is no getting around the fact that we are treating the particle as non-relativistic but the external electromagnetic field is, by it's very nature, relativistic. So with this treatment, we will destroy the Galilean invariance of the theory. The consequences of this will be apparent later on.

The canonical momentum for Lagrangian (22.1) is given by:

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = m\,\mathbf{v} + q\,\mathbf{A}(\mathbf{r},t)/c\,, \tag{22.2}$$

and the Hamiltonian is then found to be:

$$H = \mathbf{p}\cdot\mathbf{v} - L = \frac{1}{2}\,m\,\mathbf{v}^2 + q\,\phi(\mathbf{r},t) = \frac{1}{2m}\left[\,\mathbf{p} - q\,\mathbf{A}(\mathbf{r},t)/c\,\right]^2 + q\,\phi(\mathbf{r},t)\,, \tag{22.3}$$

and is the total energy. We now quantize this system using the canonical quantization procedure of Chapter 2. We let $\mathbf{r}\to\mathbf{R}$ and $\mathbf{p}\to\mathbf{P}$ become hermitian operators with commutation properties,

$$[\,X_i, P_j\,] = i\hbar\,\delta_{ij}\,,$$

with all other operators commuting. Heisenberg equations of motion can be easily found. The velocity

operator $\dot{X}_i$ is given by

$$\begin{aligned}\dot{X}_i &= [\,X_i, H\,]/i\hbar = \left[P_i - q\,A_i(\mathbf{R},t)/c\right]/m \\ \dot{P}_i &= [\,P_i, H\,]/i\hbar = \frac{m}{2}\left\{\,\dot{X}_j\frac{\partial\dot{X}_j}{\partial X_i} + \frac{\partial\dot{X}_j}{\partial X_i}\dot{X}_j\right\} - q\,\frac{\partial\phi(\mathbf{R},t)}{\partial X_i} \\ &= q\left\{\,\frac{1}{2c}\left\{\,\dot{X}_j\,\frac{\partial A_j(\mathbf{R},t)}{\partial X_i} + \frac{\partial A_j(\mathbf{R},t)}{\partial X_i}\,\dot{X}_j\,\right\} - \frac{\partial\phi(\mathbf{R},t)}{\partial X_i}\right\}.\end{aligned} \tag{22.4}$$

But from the first of (22.4), we find

$$\dot{P}_i = m\,\ddot{X}_i + \frac{q}{c}\left\{\,\frac{\partial A_i}{\partial t} + \frac{1}{2}\left\{\,\dot{X}_j\frac{\partial A_i}{\partial X_j} + \frac{\partial A_i}{\partial X_j}\dot{X}_j\,\right\}\right\}.$$

Thus we find:

$$m\,\ddot{X}_i = q\left\{\,-\frac{\partial\phi}{\partial X_i} - \frac{1}{c}\frac{\partial A_i}{\partial t} + \frac{\dot{X}_j}{2c}\left\{\,\frac{\partial A_j}{\partial X_i} - \frac{\partial A_i}{\partial X_j}\,\right\} + \left\{\,\frac{\partial A_j}{\partial X_i} - \frac{\partial A_i}{\partial X_j}\,\right\}\frac{\dot{X}_j}{2c}\,\right\},$$

or, in vector form:

$$m\,\ddot{\mathbf{R}} = q\left\{\,\mathbf{E}(\mathbf{R},t) + \left\{\,\mathbf{V}\times\mathbf{B}(\mathbf{R},t) - \mathbf{B}(\mathbf{R},t)\times\mathbf{V}\,\right\}/(2c)\,\right\}, \tag{22.5}$$

where

$$\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}\,, \qquad \mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}\,. \tag{22.6}$$

Eq. (22.5) is the quantum version of the Lorentz force on a charged particle.

Our treatment here of a charged particle in an external electromagnetic field is called "semi-classical" because we have not considered the field as part of the energy to be quantized. Thus we cannot treat problems in which the reaction of a charged particle back on the field are included.

22.1.1 Probability conservation

In the Schrödinger picture, the wave function in the coordinate basis satisfies the equation:

$$\left\{\,\frac{1}{2m}\left[\,\frac{\hbar}{i}\,\boldsymbol{\nabla} - \frac{q}{c}\,\mathbf{A}(\mathbf{r},t)\,\right]^2 + q\,\phi(\mathbf{r},t)\,\right\}\psi(\mathbf{r},t) = i\hbar\,\frac{\partial\psi(\mathbf{r},t)}{\partial t}\,. \tag{22.7}$$

The probability conservation equation obeyed by solutions of (22.7) is given by:

$$\frac{\partial\rho(\mathbf{r},t)}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{j}(\mathbf{r},t) = 0\,, \tag{22.8}$$

with the probability density and current given by:

$$\begin{aligned}\rho(\mathbf{r},t) &= |\psi(\mathbf{r},t)|^2\,, \\ \mathbf{j}(\mathbf{r},t) &= \frac{1}{2m}\left\{\,\psi^*(\mathbf{r},t)\left[\,\frac{\hbar}{i}\,\boldsymbol{\nabla} - \frac{q}{c}\,\mathbf{A}(\mathbf{r},t)\,\right]\psi(\mathbf{r},t) + \left\{\,\left[\,\frac{\hbar}{i}\,\boldsymbol{\nabla} - \frac{q}{c}\,\mathbf{A}(\mathbf{r},t)\,\right]\psi(\mathbf{r},t)\,\right\}^*\psi(\mathbf{r},t)\,\right\}, \\ &= \frac{\hbar}{2im}\left\{\,\psi^*(\mathbf{r},t)\left[\,\boldsymbol{\nabla}\psi(\mathbf{r},t)\right] - \left[\,\boldsymbol{\nabla}\psi^*(\mathbf{r},t)\right]\psi(\mathbf{r},t)\,\right\} - \frac{q}{mc}\,\mathbf{A}(\mathbf{r},t)\,\rho(\mathbf{r},t)\,.\end{aligned} \tag{22.9}$$

22.1.2 Gauge transformations

In the Heisenberg representation, the equations of motion for $\mathbf{R}(t)$, Eqs. (22.5), depend only on the electric and magnetic fields and are therefore gauge invariant. However in the Schrödinger representation, Schrödinger's equation, Eq. (22.7), depends on the potential functions, $\phi(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$, and not on the electric and magnetic fields, and so appear to be gauge *dependent*. This is, in fact, not the case, if we gauge transform the wave function as well as the fields. We shall prove the following theorem:

Theorem 53 (gauge invariance)**.** *Solutions of Schrödinger's equation are invariant under the following gauge transformation:*

$$\begin{aligned}\phi'(\mathbf{r},t) &= \phi(\mathbf{r},t) + \frac{1}{c}\frac{\partial\Lambda(\mathbf{r},t)}{\partial t}\,,\\ \mathbf{A}'(\mathbf{r},t) &= \mathbf{A}(\mathbf{r},t) - \boldsymbol{\nabla}\Lambda(\mathbf{r},t)\,,\\ \psi'(\mathbf{r},t) &= e^{iq\,\Lambda(\mathbf{r},t)/(\hbar c)}\,\psi(\mathbf{r},t)\,.\end{aligned} \tag{22.10}$$

Proof. We first compute:

$$\begin{aligned} i\hbar\,\frac{\partial\psi'(\mathbf{r},t)}{\partial t} &= e^{iq\,\Lambda(\mathbf{r},t)/(\hbar c)}\Big\{\, i\hbar\,\frac{\partial\psi(\mathbf{r},t)}{\partial t} - \frac{q}{c}\Big[\,\frac{\partial\Lambda(\mathbf{r},t)}{\partial t}\,\Big]\,\psi(\mathbf{r},t)\Big\}\,,\\ \frac{\hbar}{i}\,\boldsymbol{\nabla}\psi'(\mathbf{r},t) &= e^{iq\,\Lambda(\mathbf{r},t)/(\hbar c)}\,\Big\{\,\frac{\hbar}{i}\,\boldsymbol{\nabla}\psi(\mathbf{r},t) + \frac{q}{c}\,\big[\,\boldsymbol{\nabla}\Lambda(\mathbf{r},t)\,\big]\,\psi(\mathbf{r},t)\,\Big\}\end{aligned} \tag{22.11}$$

The last of (22.11) gives:

$$\Big\{\,\frac{\hbar}{i}\boldsymbol{\nabla} - \frac{q}{c}\mathbf{A}'(\mathbf{r},t)\,\Big\}\,\psi'(\mathbf{r},t) = e^{iq\,\Lambda(\mathbf{r},t)/(\hbar c)}\,\Big\{\,\frac{\hbar}{i}\boldsymbol{\nabla} - \frac{q}{c}\,\mathbf{A}(\mathbf{r},t)\,\Big\}\,\psi(\mathbf{r},t)\,.$$

Substitution of these results into Schrödinger's equation in the prime system, gives:

$$\Big\{\,\frac{1}{2m}\,\Big[\,\frac{\hbar}{i}\,\boldsymbol{\nabla} - \frac{q}{c}\,\mathbf{A}'(\mathbf{r},t)\,\Big]^2 + q\,\phi'(\mathbf{r},t)\,\Big\}\,\psi'(\mathbf{r},t) = i\hbar\,\frac{\partial\psi'(\mathbf{r},t)}{\partial t}\,.$$

where $\phi'(\mathbf{r},t)$ and $\mathbf{A}'(\mathbf{r},t)$ are the scalar and vector potentials in the prime system. The fact that the gauge potential $\Lambda(\mathbf{r},t)$, and thus the phase of the wave function in the transformed system, can depend on $\mathbf{r}$ and t can have significant physical consequences, which we will study further in Section 22.5. □

The probability $\rho(\mathbf{r},t)$ and the probability current density $\mathbf{j}(\mathbf{r},t)$ are invariant under a gauge transformation:

$$|\,\psi'(\mathbf{r},t)\,|^2 = |\,\psi(\mathbf{r},t)\,|^2\,, \qquad \text{and} \qquad \mathbf{j}'(\mathbf{r},t) = \mathbf{j}(\mathbf{r},t)\,. \tag{22.12}$$

22.2 Free particle in a constant electric field

In this example, we find the motion of an electron in a constant electric field. That is, we put $\mathbf{E}(\mathbf{r},t) = \mathbf{E}_0$ and set $\mathbf{B}(\mathbf{r},t) = 0$. In the Heisenberg representation, we have the equation of motion for $\mathbf{R}(t)$:

$$m\,\ddot{\mathbf{R}} = q\,\mathbf{E}_0\,.$$

So the solution is:

$$\mathbf{R}(t) = \mathbf{R}(0) + \mathbf{P}(0)\,t/m + \frac{q}{2m}\,\mathbf{E}_0\,t^2\,, \qquad \mathbf{P}(t) = \mathbf{P}(0) + \frac{q}{m}\,\mathbf{E}_0\,t\,.$$

Note that the last term on the right side of this equation is a c-number, and commutes with all operators. Thus the motion of the average value of the position of the particle is accelerated by the field, and follows the classical motion, as it must.

Let us look at the case when the electric field is in the x-direction, $\mathbf{E}_0 = E_0\,\hat{\mathbf{e}}_z$. Now since:

$$[\,X(t), X(0)\,] = -\frac{i\hbar t}{m}\,,$$

the width of a minimum wave packet in the x-direction grows linearly with time,

$$\Delta x(t) \geq \frac{\hbar t}{2m\Delta x(0)}\,,$$

like a free particle, independent of E_0.

In the Schrödinger picture, we must choose a gauge to solve the problem. We will consider two gauges. In the first gauge, we take the scalar and vector potentials to be:

$$\begin{aligned} \phi(\mathbf{r},t) &= -\mathbf{r}\cdot\mathbf{E}_0\,, \\ \mathbf{A}(\mathbf{r},t) &= 0\,. \end{aligned} \tag{22.13}$$

In the second gauge, we choose:

$$\begin{aligned} \phi'(\mathbf{r},t) &= 0\,, \\ \mathbf{A}'(\mathbf{r},t) &= -c\,\mathbf{E}_0\,t\,. \end{aligned} \tag{22.14}$$

The two gauges are connected by a gauge transformation of the form given in Eqs. (22.10) with a gauge potential given

$$\Lambda(\mathbf{r},t) = c\,(\mathbf{r}\cdot\mathbf{E}_0)\,t\,. \tag{22.15}$$

The Schrödinger equations for the two gauges are given by:

$$\left\{-\frac{\hbar^2}{2m}\,\nabla^2 - q\,\mathbf{r}\cdot\mathbf{E}_0\right\}\psi(\mathbf{r},t) = i\hbar\,\frac{\partial\psi(\mathbf{r},t)}{\partial t}\,, \tag{22.16}$$

$$\left\{\frac{1}{2m}\left[\frac{\hbar}{i}\,\boldsymbol{\nabla} + q\,\mathbf{E}_0\,t\right]^2\right\}\psi'(\mathbf{r},t) = i\hbar\,\frac{\partial\psi'(\mathbf{r},t)}{\partial t}\,. \tag{22.17}$$

The two solutions are connected by the gauge transformation,

$$\psi'(\mathbf{r},t) = e^{i\,q\,(\mathbf{r}\cdot\mathbf{E}_0)t/\hbar}\,\psi(\mathbf{r},t)\,. \tag{22.18}$$

It is easy to show that if (22.18) is substituted into (22.17) there results Eq. (22.16). We choose to solve (22.16). Take $\mathbf{E}_0$ to be in the x-direction, and consider only one dimension. Then (22.16) becomes:

$$\left\{-\frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial x^2} - q\,E_0\,x\right\}\psi(x,t) = i\hbar\,\frac{\partial}{\partial t}\psi(x,t)\,.$$

Separating variables by writing

$$\psi(x,t) = \psi_\omega(x)e^{-i\omega t}\,,$$

we find that we need to solve the equation:

$$\left\{-\frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial x^2} - q\,E_0\,x\right\}\psi_\omega(x) = \hbar\omega\psi_\omega(x)\,,$$

or

$$\left\{\frac{\partial^2}{\partial x^2} + \frac{2m\,q\,E_0}{\hbar^2}\,x + \frac{2m\,\omega}{\hbar}\right\}\psi_\omega(x) = 0\,.$$

The substitution, $x = \alpha\,\xi + x_0$, leads to the differential equation,

$$\left\{\frac{\partial^2}{\partial\xi^2} + \xi\right\}\psi_\omega(\xi) = 0\,, \tag{22.19}$$

provided we set:

$$\alpha = \left[\frac{\hbar^2}{2mqE_0}\right]^{1/3} \equiv \frac{1}{\gamma}\,, \qquad \text{and} \qquad x_0 = -\frac{\hbar\,\omega}{q\,E_0}\,. \tag{22.20}$$

Solving for ξ, we have $\xi = \gamma\,(x - x_0)$. The solutions of (22.19) are the Airy functions. From Abramowitz and Stegun [1][p. 446], solutions with the proper boundary conditions are

$$\psi_\omega(\xi) = C(\omega)\,\mathrm{Ai}(-\xi)\,,$$

where $C(\omega)$ is a constant. So the most general wave function is given by the integral:

$$\psi(x,t) = \int_{-\infty}^{+\infty} \mathrm{d}\omega\, C(\omega)\, \mathrm{Ai}\big(\,\gamma(x_0 - x)\,\big)\, e^{-i\omega t} = -\frac{qE_0}{\hbar} \int_{-\infty}^{+\infty} \mathrm{d}x_0\, C(x_0)\, \mathrm{Ai}\big(\,\gamma(x_0 - x)\,\big)\, e^{i\,q\,E_0\,x_0\,t/\hbar}\,. \quad (22.21)$$

The initial values are set by inverting this expression at $t = 0$, and solving for $C(x_0)$. This inversion, of course, is not easy to do!

22.3 The hydrogen atom

One of the main problems in physics in the early 20$^{\text{th}}$ century was to explain the structure of the hydrogen atom. The successful theoretical explanation of the energy levels and spectra of hydrogen was one of major achievements of non-relativistic quantum mechanics, and later on, of relativistic quantum mechanics, and we will study it in great detail in the following sections. We choose to first study the hydrogen atom because it is the simplest atomic system, consisting of an electron in the static coulomb potential of the proton, and illustrates many details of atomic structure. It is also satisfying to be able to calculate energy levels and eigenvectors exactly without the use of numerical methods.

An important reference work on the hydrogen atom, containing many useful results, is by Bethe and Salpeter, based on a 1932 article by Bethe in *Handbuch der Physik*, and republished as a book by Spinger-Verlag-Academic Press [2] in 1957. One can find there the standard solution of the hydrogen atom in the coordinate representation, obtaining hypergeometric functions, and first done by Schrödinger himself in 1926 [3]. We will choose a different method here, first done by Pauli, also in 1926 [4], which better illustrates the underlying symmetry of hydrogen. We explain Pauli's method in the next section.

22.3.1 Eigenvalues and eigenvectors

The Hamiltonian for the electron in the hydrogen atom center of mass system is given by:

$$H = \frac{P^2}{2m} - \frac{e^2}{R}\,, \quad (22.22)$$

where R and P are the position and momentum operators of the electron in the center of mass system of the atom. Here m is the reduced mass of the electron and e is the charge of the electron in electrostatic units. We first select units appropriate to the size of atoms — this system of units is called **atomic units**. The fine structure constant α, the Bohr radius a, the atomic unit of energy E_0, and the atomic unit of time t_0 is defined by:[1]

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137}\,, \qquad a = \frac{\hbar^2}{m\,e^2} = \frac{\hbar c}{mc^2\,\alpha} = 5.29\,\mathrm{nm} = 0.529\,\text{Å}\,, \quad (22.23)$$

$$E_0 = \frac{me^4}{\hbar^2} = \frac{e^2}{a} = mc^2\,\alpha^2 = 2 \times 13.61\,\mathrm{eV}\,, \qquad t_0 = \frac{\hbar}{E_0} = \frac{\hbar^3}{me^4} = 2.419 \times 10^{-17}\,\mathrm{s}\,.$$

It is useful to further note that the velocity of the electron in the first Bohr orbit is $v_0 = e^2/\hbar = \alpha\, c = 2.188 \times 10^6$ m/s., which is less that the velocity of light by a factor of α. This means that we should be able to use non-relativistic physics for hydrogen and take into account relativity as a small perturbation. The orbital period of the electron in the first Bohr orbit is $2\pi t_0$. In atomic units, we measure lengths in units of a, momentum in units of $\hbar/a$, angular momentum in units of $\hbar$, energy in units of E_0, and time in units of t_0. Then we can define dimensionless "barred" quantities by the following:

$$\bar{\mathbf{R}} = \mathbf{R}/a\,, \qquad \bar{\mathbf{P}} = a\,\mathbf{P}/\hbar\,, \qquad \bar{\mathbf{L}} = \mathbf{L}/\hbar \qquad \bar{H} = H/E_0\,, \qquad \bar{t} = t/t_0\,, \quad (22.24)$$

[1] We have introduced here c, the velocity of light, to correspond with the usual definitions, but nothing in our results in this section depend on c.

so that

$$[\,\bar{X}_i,\bar{P}_j\,]=i\,\delta_{i,j}\,,\qquad [\,\bar{L}_i,\bar{L}_j\,]=i\epsilon_{ijk}\,L_k\,,\qquad \frac{\mathrm{d}\bar{X}_i}{\mathrm{d}\bar{t}}=[\,\bar{X}_i,\bar{H}\,]/i\,.\qquad \frac{\mathrm{d}\bar{P}_i}{\mathrm{d}\bar{t}}=[\,\bar{P}_i,\bar{H}\,]/i\,. \tag{22.25}$$

The Rydberg unit of energy E_R is the ionization energy of hydrogen and is *one-half* the atomic unit of energy. It can be expressed conveniently in units of volts, Hertz, or nanometers by means of the formula:

$$E_R=\frac{1}{2}\,E_0=\frac{1}{2}\,mc^2\,\alpha^2=e\,V_R=2\pi\,\hbar\,f_R=2\pi\,\hbar\,c/\lambda_R\,, \tag{22.26}$$

where $V_R\approx 13.61$ eV, $f_R\approx 3.290\times 10^9$ MHz, $\lambda_R\approx 91.13$ nm, and $1/\lambda_R=109,700\,\mathrm{cm}^{-1}$.

In atomic units, the Hamiltonian $\bar{H}$ becomes:

$$\bar{H}=\frac{\bar{P}^2}{2}-\frac{1}{\bar{R}}\,. \tag{22.27}$$

with the commutation relations: $[\,\bar{X}_i,\bar{P}_j\,]=i\,\delta_{ij}$. We have now completely removed all units from the problem, so let us revert back to using unbarred coordinates and momenta, and assume that throughout we have scaled our problem using atomic units. So then in atomic units, we write:

$$H=\frac{P^2}{2}-\frac{1}{R}\,,\qquad \text{with}\qquad [\,X_i,P_j\,]=i\delta_{ij}\,. \tag{22.28}$$

The angular momentum operator, in atomic units, is then given by:

$$\mathbf{L}=\mathbf{R}\times\mathbf{P}\,, \tag{22.29}$$

and is conserved: $\dot{\mathbf{L}}=[\,\mathbf{L},H\,]/i=0$.

Exercise 72. Prove that for the Hamiltonian given by Eq. (22.28) in atomic units, that $\dot{\mathbf{L}}=0$.

For the hydrogen Hamiltonian, there is a second vector, called the Runge-Lenz vector [5, 6], that is conserved. Classically it is defined by:

$$\mathbf{A}=\mathbf{P}\times\mathbf{L}-\frac{\mathbf{R}}{R}\,. \tag{22.30}$$

Exercise 73. Prove that for the Hamiltonian given by Eq. (22.28) that the *classical* Runge-Lenz vector given in Eq. (22.30) is conserved: $\dot{\mathbf{A}}=0$. Further, show that $\mathbf{A}$ lies in the plane of the orbit so that $\mathbf{A}\cdot\mathbf{L}=0$ for all time t, and points in the direction of the perihelion of the orbit of the classical electron. For this, you will need to solve for the classical orbit equation in these units. Show also that $\mathbf{A}$ has a magnitude of $A=\epsilon=\sqrt{1+2EL^2}$, where E is the energy and ϵ the eccentricity of the orbit. (See, for example, Barger and Olsson [7][p. 144], or Wikipedia.)

In quantum mechanics, Eq. (22.30) is not a Hermitian operator and cannot be observed since $\mathbf{P}$ does not commute with $\mathbf{L}$. So we construct a Hermitian operator by adding the complex conjugate of Eq. (22.30) to Eq. (22.30) and dividing by two. This gives a quantum mechanical version of the Runge-Lenz vector:

$$\begin{aligned}\mathbf{A}&=\frac{1}{2}\,[\,\mathbf{P}\times\mathbf{L}-\mathbf{L}\times\mathbf{P}\,]-\frac{\mathbf{R}}{R}\,,\\&=\mathbf{P}\times\mathbf{L}-i\,\mathbf{P}-\frac{\mathbf{R}}{R}\,,\\&=\mathbf{R}\,P^2-\mathbf{P}\,(\,\mathbf{R}\cdot\mathbf{P}\,)-\frac{\mathbf{R}}{R}\,.\end{aligned} \tag{22.31}$$

This vector now is Hermitian and has eigenvalues and eigenvectors associated with it. The next theorem shows that the quantum mechanical Runge-Lenz vector is conserved.

Theorem 54 (Conservation of the Runge-Lenz vector)**.** *The Runge-Lenz vector* $\mathbf{A}$ *defined by Eq.* (22.31) *is conserved:* $\dot{\mathbf{A}} = 0$.

Proof. The time derivative of the Runge-Lenz vector in quantum mechanics is given by:

$$\begin{aligned}\dot{\mathbf{A}} &= [\,\mathbf{A}, H\,]/i \\ &= \frac{1}{2i}\left\{\,[\,\mathbf{P}, H\,]\times\mathbf{L} - \mathbf{L}\times[\,\mathbf{P}, H\,] - [\,\mathbf{R}, H\,]\,\frac{1}{R} - \frac{1}{R}\,[\,\mathbf{R}, H\,] - [\,1/R, H\,]\,\mathbf{R} - \mathbf{R}\,[\,1/R, H\,]\,\right\},\end{aligned} \tag{22.32}$$

since $[\,\mathbf{L}, H\,]/i = 0$. Here we have symmetrized the ordering of the operators $\mathbf{R}/R$. We first note that

$$[\,\mathbf{R}, H\,]/i = \mathbf{P}\,, \qquad \text{and} \qquad [\,\mathbf{P}, H\,]/i = -[\,\mathbf{P}, 1/R\,]/i = -\mathbf{R}/R^3\,. \tag{22.33}$$

Also

$$\begin{aligned}[\,1/R, H\,]/i &= -\frac{1}{2}\,[\,P^2, 1/R\,]/i = -\frac{1}{2}\left\{\,\mathbf{P}\cdot[\,\mathbf{P}, 1/R\,]/i + [\,\mathbf{P}, 1/R\,]/i\cdot\mathbf{P}\,\right\} \\ &= -\frac{1}{2}\left\{\,\mathbf{P}\cdot\frac{\mathbf{R}}{R^3} + \frac{\mathbf{R}}{R^3}\cdot\mathbf{P}\,\right\}.\end{aligned} \tag{22.34}$$

Substituting these results in (22.32), and working out the cross products, gives:

$$\begin{aligned}\dot{\mathbf{A}} &= \frac{1}{2}\left\{\,-\frac{\mathbf{R}}{R^3}\times(\mathbf{R}\times\mathbf{P}) + (\mathbf{R}\times\mathbf{P})\times\frac{\mathbf{R}}{R^3} - \mathbf{P}\,\frac{1}{R} - \frac{1}{R}\,\mathbf{P}\right. \\ &\qquad \left. + \frac{1}{2}\left\{\,\left(\mathbf{P}\cdot\frac{\mathbf{R}}{R^3}\right)\mathbf{R} + \left(\frac{\mathbf{R}}{R^3}\cdot\mathbf{P}\right)\mathbf{R} + \mathbf{R}\left(\mathbf{P}\cdot\frac{\mathbf{R}}{R^3}\right) + \mathbf{R}\left(\frac{\mathbf{R}}{R^3}\cdot\mathbf{P}\right)\right\}\right\} \\ &= \frac{1}{2}\left\{\,-\mathbf{R}\left(\frac{\mathbf{R}}{R^3}\cdot\mathbf{P}\right) + \frac{1}{R}\,\mathbf{P} + \mathbf{P}\,\frac{1}{R} - \mathbf{R}\left(\mathbf{P}\cdot\frac{\mathbf{R}}{R^3}\right) + i\,\frac{\mathbf{R}}{R^3} - \mathbf{P}\,\frac{1}{R} - \frac{1}{R}\,\mathbf{P}\right. \\ &\qquad \left. + \mathbf{R}\left(\frac{\mathbf{R}}{R^3}\cdot\mathbf{P}\right) + \mathbf{R}\left(\mathbf{P}\cdot\frac{\mathbf{R}}{R^3}\right) - i\,\frac{\mathbf{R}}{R^3}\,\right\} = 0\,,\end{aligned} \tag{22.35}$$

as stated in the theorem. □

Theorem 55 (Properties of the Runge-Lenz vector)**.** *The quantum mechanical Runge-Lenz vector* $\mathbf{A}$ *defined by Eq.* (22.31) *satisfies the following relations:*

$$A^2 = 1 + 2\,H\,(L^2 + 1) \equiv \epsilon^2\,, \tag{22.36}$$

and

$$\mathbf{A}\cdot\mathbf{L} = \mathbf{L}\cdot\mathbf{A} = 0\,. \tag{22.37}$$

Proof. The proof is left as an exercise. Note that the quantum mechanical definition of the eccentricity ϵ differs slightly from the classical definition. □

Theorem 56 (Commutation relations of the Runge-Lenz vector)**.** *The commutation relations of the Runge-Lenz vector with the angular moment vector are given by:*

$$[\,A_i, A_j\,] = -2\,i\,H\,\epsilon_{ijk}\,L_k\,, \qquad [\,A_i, L_j\,] = i\,\epsilon_{ijk}A_k\,, \qquad [\,L_i, L_j\,] = i\,\epsilon_{ijk}L_k\,. \tag{22.38}$$

Proof. We also leave this proof as an exercise for the reader. Take our word for it! □

If the spectra of H is negative definite, then we see from Eqs. (22.38) that if we define a new vector $\mathbf{K}$ by:

$$\mathbf{K} = \mathbf{A}\,N = N\,\mathbf{A}\,, \qquad \text{where} \qquad N = \frac{1}{\sqrt{-2H}}\,. \tag{22.39}$$

Recall that H commutes with $\mathbf{A}$. Then Eqs. (22.38) become:

$$[\,K_i, K_j\,] = i\epsilon_{ijk}\,L_k\,, \qquad [\,K_i, L_j\,] = i\,\epsilon_{ijk}K_k\,, \qquad [\,L_i, L_j\,] = i\,\epsilon_{ijk}L_k\,, \tag{22.40}$$

with $\mathbf{K}\cdot\mathbf{L} = \mathbf{L}\cdot\mathbf{K} = 0$. Furthermore Eq. (22.36) becomes:

$$K^2 + L^2 + 1 = N^2 = -1/(2H)\,, \qquad H = -\frac{1}{2\,(K^2 + L^2 + 1)}\,. \tag{22.41}$$

In order to uncouple the system described by (22.40), we define two new vector $\mathbf{J}_1$ and $\mathbf{J}_2$ by:

$$\mathbf{J}_1 = (\,\mathbf{L} + \mathbf{K}\,)/2\,, \qquad \mathbf{J}_2 = (\,\mathbf{L} - \mathbf{K}\,)/2\,. \tag{22.42}$$

Then we find the relations:

$$K^2 + L^2 = 2\,(\,J_1^2 + J_2^2\,)\,, \qquad \mathbf{K}\cdot\mathbf{L} = J_1^2 - J_2^2 = 0\,. \tag{22.43}$$

and the commutation relations:

$$[\,J_{1,i}, J_{2,j}\,] = i\epsilon_{ijk}\,J_{1,k}\,, \qquad [\,J_{2,i}, J_{2,j}\,] = i\,\epsilon_{ijk}J_{2,k}\,, \qquad [\,J_{1,i}, J_{2,j}\,] = 0\,. \tag{22.44}$$

So $\mathbf{J}_1$ and $\mathbf{J}_2$ are two commuting operators which obey the commutation relations of angular momentum, and therefore have common direct product eigenvectors which we define by:

$$\begin{aligned} J_1^2\,|\,j_1, m_1, j_2, m_2\,\rangle &= j_1(j_1+1)\,|\,j_1, m_1, j_2, m_2\,\rangle\,, \qquad J_{1,z}\,|\,j_1, m_1, j_2, m_2\,\rangle = m_1\,|\,j_1, m_1, j_2, m_2\,\rangle\,, \\ J_2^2\,|\,j_1, m_1, j_2, m_2\,\rangle &= j_2(j_2+1)\,|\,j_1, m_1, j_2, m_2\,\rangle\,, \qquad J_{2,z}\,|\,j_1, m_1, j_2, m_2\,\rangle = m_2\,|\,j_1, m_1, j_2, m_2\,\rangle\,, \end{aligned}$$

with j_1 and j_2 given by $0, 1/2, 1, 3/2, 2, \ldots$ and $-j_1 \le m_1 \le +j_1$ and $-j_2 \le m_2 \le +j_2$. However since $J_1^2 = J_2^2$, the only eigenvectors allowed are those for which $j_1 = j_2 \equiv j$. So let us set $j = j_1 = j_2 = (n-1)/2$, with $n = 0, 1, 2, \ldots$. Furthermore, from Eq. (22.42), we see that the physical angular momentum vector $\mathbf{L}$ is given by the *sum* of $\mathbf{J}_1$ and $\mathbf{J}_2$:

$$\mathbf{L} = \mathbf{J}_1 + \mathbf{J}_2\,. \tag{22.45}$$

So the *coupled* angular momentum state $|\,(j_1, j_2)\,\ell, m_\ell\,\rangle$ given by:

$$|\,(j_1, j_2)\,\ell, m\,\rangle = \sum_{m_1, m_2} \langle\,j_1, m_1, j_2, m_2\,|\,(j_1, j_2)\,\ell, m\,\rangle\,|\,j_1, m_1, j_2, m_2\,\rangle\,, \tag{22.46}$$

where the bracket is a Clebsch-Gordan coefficient, is an eigenvector of J_1^2, J_2^2, L^2, and L_z (See Section 21.4.1). If we set $j_1 = j_2 = j$, this is just the state that we want. So let us put:

$$|\,n, \ell, m\,\rangle \equiv |\,(j, j)\,\ell, m\,\rangle = \sum_{m_1, m_2} \langle\,j, m_1, j, m_2\,|\,(j, j)\,\ell, m\,\rangle\,|\,j, m_1, j, m_2\,\rangle\,, \tag{22.47}$$

where $j = (n-1)/2$, and from the triangle inequality for coupled states, $\ell = 0, 1, \ldots, (n-1)$, with $-\ell \le m = m_1 + m_2 \le +\ell$. Now from Eqs. (22.41) and (22.43), the Hamilton is given by:

$$H = -\frac{1}{2\,\big(\,2(\,J_1^2 + J_2^2\,) + 1\,\big)} = -\frac{1}{2\,\big(4J_1^2 + 1\big)}\,, \tag{22.48}$$

the eigenvector defined in Eq. (22.47) is also an eigenvector of H:

$$\begin{aligned} &H\,|\,n, \ell, m_\ell\,\rangle = E_n\,|\,n, \ell, m_\ell\,\rangle\,, \\ &\text{with} \qquad E_n = -\frac{1}{2\,\big(\,2j(j+1) + 1\,\big)} = -\frac{1}{2\,\big(\,(n-1)(n+1) + 1\,\big)} = -\frac{1}{2\,n^2}\,. \end{aligned} \tag{22.49}$$

From the definition of N is Eq. (22.39), we find that N is diagonal in eigenvectors of the Hamiltonian:

$$N\,|\,n,\ell,m_\ell\,\rangle = n\,|\,n,\ell,m_\ell\,\rangle\,. \tag{22.50}$$

For a fixed value of n, there are $(2j+1)^2 = n^2$ values of m_1 and m_2, so the degeneracy of the n^{th} state is n^2. The unusual degeneracy here is a result of the conserved Lenz vector for the coulomb potential. Our method of writing the Hamiltonian in terms of the Lenz and angular momentum vectors has reduced the eigenvalue problem for the hydrogen atom to an algebra.

Exercise 74. Work out the eigenvectors for the $n = 1$, 2, and 3 levels of hydrogen in both the $|\,j,m_1,j,m_2\,\rangle$ representation and in the coupled representation $|\,(j,j)\,\ell,m\,\rangle$. Use the results in Table 21.1 for the Clebsch-Gordan coefficients.

22.3.2 Matrix elements of the Runge-Lenz vector

Matrix elements of the Runge-Lenz vector $\mathbf{K} = N\,\mathbf{A}$ in eigenvectors of the hydrogen atom can be found easily by using the Wigner-Eckart theorem and the angular momentum theorems of Section 21.5.3. Writing the vectors as tensor operators of rank one, we first compute reduced matrix elements of $\mathbf{J}_1$ and $\mathbf{J}_2$. We find:

$$\begin{aligned}\langle\,(j,j)\,\ell\,\|\,J_1\,\|\,(j',j')\,\ell'\,\rangle &= \delta_{j,j'}\,(-)^{2j+\ell'+1}\,\sqrt{(2\ell+1)(2\ell'+1)}\begin{Bmatrix} j & \ell & j \\ \ell' & j & 1\end{Bmatrix}\langle\,j\,\|\,J_1\,\|\,j\,\rangle \\ &= \delta_{j,j'}\,(-)^{2j+\ell'+1}\,\sqrt{(2\ell+1)(2\ell'+1)}\begin{Bmatrix} j & \ell & j \\ \ell' & j & 1\end{Bmatrix}\sqrt{2j(2j+1)(2j+2)}/2\end{aligned} \tag{22.51}$$

and

$$\begin{aligned}\langle\,(j,j)\,\ell\,\|\,J_2\,\|\,(j',j')\,\ell'\,\rangle &= \delta_{j,j'}\,(-)^{2j+\ell+1}\,\sqrt{(2\ell+1)(2\ell'+1)}\begin{Bmatrix} j & \ell & j \\ \ell' & j & 1\end{Bmatrix}\langle\,j\,\|\,J_2\,\|\,j\,\rangle \\ &= \delta_{j,j'}\,(-)^{2j+\ell+1}\,\sqrt{(2\ell+1)(2\ell'+1)}\begin{Bmatrix} j & \ell & j \\ \ell' & j & 1\end{Bmatrix}\sqrt{2j(2j+1)(2j+2)}/2\end{aligned} \tag{22.52}$$

Now $\mathbf{K} = \mathbf{J}_1 - \mathbf{J}_2$, so

$$\begin{aligned}&\langle\,(j,j)\,\ell\,\|\,K\,\|\,(j',j')\,\ell'\,\rangle \\ &\qquad = \delta_{j,j'}\,(-)^{2j+1}\,\frac{\left[\,(-)^{\ell'}-(-)^{\ell}\,\right]}{2}\,\sqrt{(2\ell+1)(2\ell'+1)2j(2j+1)(2j+2)}\begin{Bmatrix} j & \ell & j \\ \ell' & j & 1\end{Bmatrix}.\end{aligned} \tag{22.53}$$

Now the $6j$-symbol vanishes unless $\ell' = \ell-1, \ell, \ell+1$, but the factor $\left[\,(-)^{\ell'}-(-)^{\ell}\,\right]$ vanishes for $\ell' = \ell$. So, recalling that $2j = n-1$, and switching the notation to $|\,n,\ell,m\,\rangle$, the only non-vanishing reduced matrix elements are:

$$\begin{aligned}\langle\,n,\ell\,\|\,K\,\|\,n,\ell-1\,\rangle &= \sqrt{\ell\,(n^2-\ell^2)}\,, \\ \langle\,n,\ell\,\|\,K\,\|\,n,\ell+1\,\rangle &= -\sqrt{(\ell+1)\,(n^2-(\ell+1)^2)}\,.\end{aligned} \tag{22.54}$$

So matrix elements of the Runge-Lenz vector connect eigenvectors of the hydrogen Hamiltonian with the same values of the principle quantum number n but values of $\ell' = \ell \pm 1$. That is, the Runge-Lenz vector is a ladder operator for the total angular momentum quantum number within the same principle quantum number.

Since $\mathbf{L} = \mathbf{J}_1 + \mathbf{J}_2$, we can easily check that

$$\langle\,n,\ell\,\|\,L\,\|\,n,\ell'\,\rangle = \delta_{\ell,\ell'}\,\sqrt{2\ell\,(2\ell+1)\,(2\ell+2)}/2\,, \tag{22.55}$$

in agreement with (21.283).

Exercise 75. Using Eqs. (22.51) and (22.52), prove Eq. (22.55).

22.3.3 Symmetry group

The symmetry group associated with the coulomb potential is apparently $SO(3) \times SO(3) \sim SO(4)$. So let us review the properties of the $SO(4)$ group. Consider orthogonal transformations in four Euclidean dimensions, given by:

$$x'_\alpha = R_{\alpha,\beta}\, x_\beta \,, \tag{22.56}$$

where the Greek indices α, β run from 0 to 3, and where R orthogonal: $R^T R = 1$. The set of all such matrices form the $O(4)$ group. Infinitesimal transformations are given by: $R_{\alpha,\beta} = \delta_{\alpha,\beta} + \Delta\omega_{\alpha,\beta} + \cdots$. Infinitesimal unitary transformations in Hilbert space are given by:

$$U(1 + \Delta\omega) = 1 + \frac{i}{\hbar}\left\{ \frac{1}{2}\,\Delta\omega_{\alpha,\beta}\, J_{\alpha,\beta} + \cdots \right\}, \tag{22.57}$$

where $J_{\alpha,\beta} = -J_{\beta,\alpha}$ are the generators of the group. Since $J_{\alpha,\beta}$ are antisymmetric, there are six of them. The generators transform according to:

$$U^\dagger(R)\, J_{\alpha,\beta}\, U(R) = R_{\alpha,\alpha'}\, R_{\beta,\beta'}\, J_{\alpha',\beta'} \,. \tag{22.58}$$

Setting $R = 1 + \Delta\omega$, we find that the generators obey the algebra:

$$[\, J_{\alpha,\beta}, J_{\alpha',\beta'}\,] = i\hbar\left\{ \delta_{\alpha,\alpha'}\, J_{\beta,\beta'} + \delta_{\beta,\beta'}\, J_{\alpha,\alpha'} - \delta_{\alpha,\beta'}\, J_{\beta,\alpha'} - \delta_{\beta,\alpha'}\, J_{\alpha,\beta'} \right\}. \tag{22.59}$$

One can easily check that a realization of this algebra is obtained by setting:

$$J_{\alpha,\beta} = X_\alpha\, P_\beta - X_\beta\, P_\alpha \,, \tag{22.60}$$

where X_α and P_β obey the commutation rules:

$$[\, X_\alpha, P_\beta\,] = i\hbar\,\delta_{\alpha,\beta} \,, \qquad [\, X_\alpha, X_\beta\,] = [\, P_\alpha, P_\beta\,] = 0 \,. \tag{22.61}$$

This realization of $J_{\alpha,\beta}$ is a generalization of angular momentum in four Euclidean dimensions.

Exercise 76. Show that $J_{\alpha,\beta}$ defined by Eq. (22.60) with X_α and P_β satisfying (22.61), satisfies Eq. (22.59).

However we can also satisfy the commutation rules for the generators of $SO(4)$ by embedding the six operators $\mathbf{K}$ and $\mathbf{L}$ in $J_{\alpha,\beta}$ with the definitions:

$$J_{0,i} = -J_{i,0} = K_i \,, \qquad \text{and} \qquad J_{i,j} = \epsilon_{ijk}\, L_k \,, \tag{22.62}$$

where the Roman indices i, j, k run from 1 to 3. Explicitly, $J_{\alpha,\beta}$ is given by the matrix:

$$J_{\alpha,\beta} = \begin{pmatrix} 0 & K_1 & K_2 & K_3 \\ -K_1 & 0 & L_3 & -L_2 \\ -K_2 & -L_3 & 0 & L_1 \\ -K_3 & L_2 & -L_1 & 0 \end{pmatrix} . \tag{22.63}$$

Exercise 77. Using the commutation relations (22.40), show that the definitions (22.63) satisfy the algebra of Eq. (22.59).

Now for $(\alpha, \beta) = (i, j)$, the definitions (22.60) are exactly the angular momentum of the particle: X_i and P_i. That is

$$J_{i,j} = \epsilon_{ijk}\, L_k = X_i P_j - X_j P_i \,. \tag{22.64}$$

For $J_{0,i}$, we find:

$$J_{0,i} = K_i = N\, A_i = A_i\, N = X_0\, P_i - X_i\, P_0 \,. \tag{22.65}$$

It is not at all easy to find X_0 and P_0 in terms of the dynamic variables X_i and P_i, the major difficulty being the representation of N in terms of X_i and P_i.

The quantity:

$$\frac{1}{2}\, J_{\alpha,\beta} J_{\alpha,\beta} = K^2 + L^2 = N^2 - 1 \geq 0 \,, \tag{22.66}$$

is a Casmir invariant for the $SO(4)$ group.

22.3.4 Operator factorization

A completely different method of solving for the eigenvalues and eigenvectors for the hydrogen Hamiltonian is the operator factorization method, first done by Schrödinger in 1940 [8]. The method is further expanded to several kinds of potential problems in an article by Infeld and Hull [9].

In Section 21.1.3, we introduced a Hermitian radial linear momentum operator P_r defined by (here $\hbar \to 1$):

$$P_r = \frac{1}{R}\left[\,\mathbf{R}\cdot\mathbf{P} - i\,\right] \mapsto \frac{1}{i}\left[\frac{\partial}{\partial r} + \frac{1}{r}\right] = \frac{1}{i}\left[\frac{1}{r}\frac{\partial}{\partial r}\,r\right] . \tag{22.67}$$

We also showed that $P^2 = P_r^2 + L^2/R^2$. So the hydrogen Hamiltonian can be written as:

$$H = \frac{P_r^2}{2} + \frac{L^2}{2R^2} - \frac{1}{R} . \tag{22.68}$$

Now let $|\,\epsilon,\ell,m\,\rangle$ be an eigenvector of L^2 and L_z with eigenvalues $\ell(\ell+1)$ and m respectively. Then for the radial equation, we write:

$$H_\ell\,|\,\epsilon,\ell\,\rangle = \epsilon\,|\,\epsilon,\ell\,\rangle\,, \qquad \text{with} \qquad H_\ell = \frac{P_r^2}{2} + \frac{\ell(\ell+1)}{2R^2} - \frac{1}{R}\,, \tag{22.69}$$

where $\langle\, r\,|\,n,\ell\,\rangle = R_{n\ell}(r)$. Now the first thing to note here is that the Hamiltonian H_ℓ depends on ℓ and so the eigenvectors $|\,\epsilon,\ell\,\rangle$ *are not orthogonal with respect to* ℓ. That is, ℓ must be regarded here as a parameter, not an eigenvalue, of H_ℓ.

The operator factorization method consists of factoring H_ℓ into two parts of the form: $H_\ell = A_\ell^\dagger\,A_\ell + c_\ell$, where c_ℓ is some constant. So let us try to find an operator A_ℓ of the form:

$$A_\ell = \frac{1}{\sqrt{2}}\left\{\,P_r + \frac{i\alpha_\ell}{R} - i\beta_\ell\,\right\}, \tag{22.70}$$

with α_ℓ and β_ℓ real numbers. Then we find:

$$\begin{aligned} 2\,A_\ell^\dagger\,A_\ell &= P_r^2 + i\alpha_\ell\,[\,P_r, \frac{1}{R}\,] + \frac{\alpha_\ell^2}{R^2} - \frac{2\alpha_\ell\beta_\ell}{R} + \beta_\ell^2 \\ &= P_r^2 + i\alpha_\ell\,\frac{1}{R}\,[\,R, P_r\,]\,\frac{1}{R} + \frac{\alpha_\ell^2}{R^2} - \frac{2\alpha_\ell\beta_\ell}{R} + \beta_\ell^2 \\ &= P_r^2 + \frac{\alpha_\ell(\alpha_\ell-1)}{R^2} - \frac{2\alpha_\ell\beta_\ell}{R} + \beta_\ell^2\,. \end{aligned} \tag{22.71}$$

So we need to require that $\alpha_\ell(\alpha_\ell-1) = \ell(\ell+1)$ and $\alpha_\ell\beta_\ell = 1$. There are two solutions to these equations: $\alpha_\ell = \ell+1$, in which case $\beta_\ell = 1/(\ell+1)$, and $\alpha_\ell = -\ell$, in which case $\beta_\ell = -1/\ell$. For the first of these solutions, we have:

$$A_\ell^{(+)} = \frac{1}{\sqrt{2}}\left\{\,P_r + \frac{i(\ell+1)}{R} - \frac{i}{(\ell+1)}\,\right\}, \qquad A_\ell^{(+)\,\dagger} = \frac{1}{\sqrt{2}}\left\{\,P_r - \frac{i(\ell+1)}{R} + \frac{i}{(\ell+1)}\,\right\}, \tag{22.72}$$

and for the second solution, we have:

$$A_\ell^{(-)} = \frac{1}{\sqrt{2}}\left\{\,P_r - \frac{i\ell}{R} + \frac{i}{\ell}\,\right\}, \qquad A_\ell^{(-)\,\dagger} = \frac{1}{\sqrt{2}}\left\{\,P_r + \frac{i\ell}{R} - \frac{i}{\ell}\,\right\}. \tag{22.73}$$

For these second solutions, we must require that $\ell \neq 0$. However, we see that these two solutions are related. That is, since P_r is Hermitian,

$$A_{\ell+1}^{(-)} = A_\ell^{(+)\,\dagger}\,, \qquad \text{and} \qquad A_{\ell+1}^{(-)\,\dagger} = A_\ell^{(+)}\,. \tag{22.74}$$

That is the (−) solutions interchange creation and annihilation operators of the (+) solutions and decrement the ℓ value by one unit. So they do not add any new information and we only need to consider the (+) set of solutions. In the following, we choose the (+) solutions and omit the (+) designation from now one. We find:

$$A_\ell^\dagger\, A_\ell = \frac{P_r^2}{2} + \frac{\ell(\ell+1)}{2R^2} - \frac{1}{R} + \frac{1}{2(\ell+1)^2} = H_\ell + \frac{1}{2(\ell+1)^2}\,, \tag{22.75}$$

$$A_\ell\, A_\ell^\dagger = \frac{P_r^2}{2} + \frac{(\ell+2)(\ell+1)}{2R^2} - \frac{1}{R} + \frac{1}{2(\ell+1)^2} = H_{\ell+1} + \frac{1}{2(\ell+1)^2}\,. \tag{22.76}$$

So setting $\ell \to \ell+1$ in (22.75) and substituting into (22.76) gives:

$$A_\ell\, A_\ell^\dagger = A_{\ell+1}^\dagger\, A_{\ell+1} + \frac{1}{2(\ell+1)^2} - \frac{1}{2(\ell+2)^2}\,. \tag{22.77}$$

Eq. (22.77) is very much like a commutation relation, except that it involves different values of ℓ for the reverse product. Nevertheless, since we have been able to factor the Hamiltonian, the eigenvalue problem for H becomes:

$$A_\ell^\dagger\, A_\ell\,|\,\epsilon,\ell\,\rangle = \left\{\,\epsilon + \frac{1}{2\,(\ell+1)^2}\,\right\}|\,\epsilon,\ell\,\rangle\,. \tag{22.78}$$

The left-hand-side of this expression is positive definite. Multiplying on the left by $\langle\,\epsilon,\ell\,|$ gives:

$$|\,A_\ell\,|\,\epsilon,\ell\,\rangle\,|^2 = \epsilon + \frac{1}{2\,(\ell+1)^2} \geq 0\,, \tag{22.79}$$

which means that for fixed $\epsilon < 0$, $0 \leq \ell \leq 1/\sqrt{-2\epsilon} - 1$. Operating on Eq. (22.77), by A_ℓ gives

$$A_\ell\, A_\ell^\dagger\, A_\ell\,|\,\epsilon,\ell\,\rangle = \left\{\,\epsilon + \frac{1}{2\,(\ell+1)^2}\,\right\} A_\ell\,|\,\epsilon,\ell\,\rangle\,, \tag{22.80}$$

and using (22.77) we find:

$$\left\{\,A_{\ell+1}^\dagger\, A_{\ell+1} + \frac{1}{2(\ell+1)^2} - \frac{1}{2(\ell+2)^2}\,\right\} A_\ell\,|\,\epsilon,\ell\,\rangle = \left\{\,\epsilon + \frac{1}{2\,(\ell+1)^2}\,\right\} A_\ell\,|\,\epsilon,\ell\,\rangle\,, \tag{22.81}$$

or

$$\left\{\,A_{\ell+1}^\dagger\, A_{\ell+1}\right\} A_\ell\,|\,\epsilon,\ell\,\rangle = \left\{\,\epsilon + \frac{1}{2\,(\ell+2)^2}\,\right\} A_\ell\,|\,\epsilon,\ell\,\rangle\,, \tag{22.82}$$

and comparison with Eq. (22.78) gives

$$A_\ell\,|\,\epsilon,\ell\,\rangle = c_{\epsilon,\ell+1}\,|\,\epsilon,\ell+1\,\rangle\,, \tag{22.83}$$

where $c_{\epsilon,\ell+1}$ is some constant. So A_ℓ, when operating on $|\,\epsilon,\ell\,\rangle$, *increases* the value of ℓ by one with the same value of ϵ. Since ℓ is bounded from above by $\ell \leq 1/\sqrt{-2\epsilon} - 1$, this can continue only until for some value $\ell = \ell_{\max}$, the right-hand side of (22.83) gives zero. That is:

$$A_{\ell_{\max}}\,|\,\epsilon,\ell_{\max}\,\rangle = 0\,. \tag{22.84}$$

Operating on (22.84) by $A_\ell^\dagger$ gives:

$$A_{\ell_{\max}}^\dagger\, A_{\ell_{\max}}\,|\,\epsilon,\ell_{\max}\,\rangle = 0 = \left\{\,\epsilon + \frac{1}{2\,(\ell_{\max}+1)^2}\,\right\}|\,\epsilon,\ell_{\max}\,\rangle\,. \tag{22.85}$$

So let us put $n = \ell_{\max} + 1 = 1, 2, \ldots$, then (22.85) requires that

$$\epsilon = -\frac{1}{2\,(\ell_{\max}+1)^2} = -\frac{1}{2\,n^2}\,. \tag{22.86}$$

Then, for fixed n, $\ell = 0, 1, \ldots, n-1$. Eq. (22.86) is the same result we got in Eq. (22.49) using the algebraic method of Section 22.3.1.

So now label ϵ by n and put $|\,\epsilon, \ell\,\rangle \to |\,n, \ell\,\rangle$. The normalization factor $c_{\ell+1}$ in can be found by taking the inner product of Eq. (22.83) with itself. We find:

$$|\,c_{n,\ell+1}\,|^2 = \langle\, n, \ell\,|\, A_\ell^\dagger\, A_\ell\,|\, n, \ell\,\rangle = \frac{1}{2}\left\{\, \frac{1}{(\ell+1)^2} - \frac{1}{n^2}\,\right\} = \frac{n^2 - (\ell+1)^2}{2\,n^2\,(\ell+1)^2}\,, \tag{22.87}$$

so

$$c_{n,\ell} = -i\,\frac{\sqrt{[\,n^2 - \ell^2\,]/2}}{n\,\ell}\,. \tag{22.88}$$

Here the phase is chosen so that the wave functions in coordinate space are all real. Then the normalized operator which increases the value of ℓ for fixed n gives the result:

$$|\,n, \ell+1\,\rangle = \frac{i\,n\,(\ell+1)}{\sqrt{[\,n^2 - (\ell+1)^2\,]/2}}\,A_\ell\,|\,n, \ell\,\rangle\,, \qquad \text{for } \ell = 0, 1, \ldots, n-2, \text{ and } n > 1, \tag{22.89}$$

which generates the $|\,n, \ell+1\,\rangle$ radial state from the $|\,n, \ell\,\rangle$ one. For the special state when $\ell = n-1$, we have:

$$A_{n-1}\,|\,n, n-1\,\rangle = \frac{1}{\sqrt{2}}\left\{\, P_r + \frac{in}{R} - \frac{i}{n}\,\right\}|\,n, n-1\,\rangle = 0\,. \tag{22.90}$$

In the coordinate representation, this gives:

$$\left\{\, \frac{\partial}{\partial r} + \frac{1-n}{r} + \frac{1}{n}\,\right\} R_{n,n-1}(r) = 0\,, \tag{22.91}$$

which has the normalized solution:

$$R_{n,n-1}(r) = \sqrt{\frac{2^{2n+1}}{n^{2n+1}\,(2n)!}}\;\, r^{n-1}\, e^{-r/n}\,. \tag{22.92}$$

We still need to find a way to find an operator to lower the ℓ value for a fixed n. This is obtained by operating on (22.78) by $A_{\ell-1}^\dagger$. Doing this, we find:

$$A_{\ell-1}^\dagger\, A_\ell^\dagger\, A_\ell\,|\,n, \ell\,\rangle = \left\{\, -\frac{1}{2\,n^2} + \frac{1}{2\,(\ell+1)^2}\,\right\} A_{\ell-1}^\dagger\,|\,n, \ell\,\rangle\,. \tag{22.93}$$

Using (22.77) gives:

$$A_{\ell-1}^\dagger\left\{\, A_{\ell-1}\, A_{\ell-1}^\dagger - \frac{1}{2\ell^2} + \frac{1}{2(\ell+1)^2}\,\right\}|\,n, \ell\,\rangle = \left\{\, -\frac{1}{2\,n^2} + \frac{1}{2\,(\ell+1)^2}\,\right\} A_{\ell-1}^\dagger\,|\,n, \ell\,\rangle\,, \tag{22.94}$$

or

$$\left\{\, A_{\ell-1}^\dagger\, A_{\ell-1}\,\right\} A_{\ell-1}^\dagger\,|\,n, \ell\,\rangle = \left\{\, -\frac{1}{2\,n^2} + \frac{1}{2\,\ell^2}\,\right\} A_{\ell-1}^\dagger\,|\,n, \ell\,\rangle\,, \tag{22.95}$$

so

$$A_{\ell-1}^\dagger\,|\,n, \ell\,\rangle = d_{n,\ell}\,|\,n, \ell-1\,\rangle\,, \tag{22.96}$$

where $d_{n,\ell}$ is some constant. So $A_{\ell-1}^\dagger$ when operating on $|\,n, \ell\,\rangle$ *decreases* the value of ℓ for the same value of n. The normalization factor $d_{n,\ell-1}$ is again found by computing the inner product of (22.96) with itself. This gives:

$$\begin{aligned} |d_{n,\ell}|^2 &= \langle\, n, \ell\,|\, A_{\ell-1}\, A_{\ell-1}^\dagger\,|\, n, \ell\,\rangle = \langle\, n, \ell\,|\left\{\, A_\ell^\dagger\, A_\ell + \frac{1}{2\ell^2} - \frac{1}{2(\ell+1)^2}\,\right\}|\,n, \ell\,\rangle \\ &= \frac{1}{2}\left\{\, \frac{1}{\ell^2} - \frac{1}{n^2}\,\right\} = \frac{n^2 - \ell^2}{2\,n^2\,\ell^2}\,, \qquad d_{n,\ell} = -i\,\frac{\sqrt{[\,n^2 - \ell^2\,]/2}}{n\,\ell}\,. \end{aligned} \tag{22.97}$$

$$
\begin{aligned}
R_{1,0}(r) &= 2\,e^{-r}\,, \\
R_{2,1}(r) &= \frac{1}{2\sqrt{6}}\,r\,e^{-r/2}\,, \\
R_{2,0}(r) &= \frac{1}{\sqrt{2}}\left(1-\frac{1}{2}\,r\right)e^{-r/2}\,, \\
R_{3,2}(r) &= \frac{4}{81\sqrt{30}}\,r^2\,e^{-r/3}\,, \\
R_{3,1}(r) &= \frac{8}{27\sqrt{6}}\left(r-\frac{1}{6}\,r^2\right)e^{-r/3}\,, \\
R_{3,0}(r) &= \frac{2}{3\sqrt{3}}\left(1-\frac{2}{3}\,r+\frac{2}{27}\,r^2\right)e^{-r/3}\,.
\end{aligned}
$$

Table 22.1: The first few radial wave functions for hydrogen.

Again the phase is chosen so that the wave functions in coordinate space are all real. Note that $d_{n,\ell} = c_{n,\ell}$. Then the normalized operator which decreases the value of ℓ for fixed n is given by:

$$
|\,n,\ell-1\,\rangle = \frac{i\,n\,\ell}{\sqrt{[\,n^2-\ell^2\,]/2}}\,A^\dagger_{\ell-1}\,|\,n,\ell\,\rangle\,, \qquad \text{for } \ell = 1,2,\ldots,n-1, \text{ and } n>1. \tag{22.98}
$$

We can use (22.98) to operate on the state with the maximum value of $\ell = n-1$ given in Eq. (22.92) to obtain all the states for that fixed value of n. For example, for the $R_{2,0}(r)$ wave function, we have:

$$
R_{2,0}(r) = \frac{2}{\sqrt{3}}\left[\frac{\partial}{\partial r}+\frac{2}{r}-1\right]R_{2,1}(r) = \frac{1}{\sqrt{2}}\left(1-\frac{1}{2}\,r\right)e^{-r/2}\,. \tag{22.99}
$$

The general result of this process produces Laguerre polynomials for each value of the principle quantum number n. The first few radial wave functions for hydrogen are given in Table 22.1. Notice that none of the $\ell = 0$ wave functions vanish at the origin. This is because of the $1/r$ singularity of the coulomb potential.

In this section, we have been able to find step operators to generate eigenvectors for total angular momentum ℓ for the hydrogen Hamiltonian using the operator factorization method. However, we still do not know what the step operators are for the principle quantum number n. We find these operators in the next section.

22.3.5 Operators for the principle quantum number

In the last section, we found operators which generated all the eigenvectors for values of the total angular momentum quantum number ℓ for the radial solutions of the hydrogen Hamiltonian for fixed values of the principle quantum number n. In this section, we obtain operators which generate all the eigenvectors for the principle quantum number n for fixed values of the total angular momentum quantum number ℓ, just the opposite result as in the last section. We do this in a round-about way, by first writing down three operators which obey $SO(2,1)$ algebra, and then relating them to the radial eigenvectors of the last section. We follow, more or less, the development by Hecht [10].

We start by defining generators of the $SO(2,1)$ algebra and proving a theorem concerning the eigenvalues and eigenvectors of these generators.

Definition 41 (Generators of $SO(2,1)$ algebra)**.** Three Hermitian operators T_1, T_2 and T_3 are generators of the $SO(2,1)$ algebra if they satisfy the commutation rules:

$$
[\,T_1,T_2\,] = -i\,T_3\,, \qquad [\,T_2,T_3\,] = i\,T_1\,, \qquad [\,T_3,T_1\,] = i\,T_2\,. \tag{22.100}
$$

The "magnitude" T^2 is defined by:

$$
T^2 = T_3^2 - T_1^2 - T_2^2\,, \qquad \text{with} \qquad [\,T^2,T_3\,] = 0\,. \tag{22.101}
$$

We also define $T_\pm = T_1 \pm iT_2$, so that $T_\pm^\dagger = T_\mp$, and with the properties:

$$[\,T_3, T_\pm\,] = \pm T_\pm\,, \qquad [\,T_+, T_-\,] = -2\,T_3\,. \tag{22.102}$$

The operator T^2 can be written in a number of ways:

$$T^2 = T_3^2 - \frac{1}{2}\,(\,T_+T_- + T_-T_+\,) = T_3^2 - T_3 - T_+T_- = T_3^2 + T_3 - T_-T_+\,. \tag{22.103}$$

The common eigenvalues and eigenvectors of T^2 and T_3 is given in the next theorem.

Theorem 57 (eigenvalues and eigenvectors for the $SO(2,1)$ algebra)**.** *Common eigenvalues and eigenvectors of the operators T^2 and T_3, are given by:*

$$\begin{aligned} T^2\,|\,k,q\,) &= k(k+1)\,|\,k,q\,)\,,\\ T_3\,|\,k,q\,) &= q\,|\,k,q\,)\,,\\ T_\pm\,|\,k,q\,) &= A_\pm(k,q)\,|\,k,q\pm 1\,)\,, \end{aligned} \tag{22.104}$$

where

$$A_\lambda(k,q) = \begin{cases} \sqrt{(q-k)(q+k+1)}\,, & \text{for } \lambda = +1,\\ \sqrt{(q+k)(q-k-1)}\,, & \text{for } \lambda = -1. \end{cases} \tag{22.105}$$

For $k \geq 0$, $q = k+1, k+2, k+3, \ldots$, and $q = -k-1, -k-2, -k-3, \ldots$. For $k < 0$, $q = q_0, q_0 \pm 1, q_0 \pm 2, \ldots$, where q_0 is arbitrary.

Proof. Let us write the general eigenvalue problems as:

$$T^2\,|\,\lambda,q\,) = \lambda|\,\lambda,q\,)\,, \qquad T_3\,|\,\lambda,q\,) = q\,|\,\lambda,q\,)\,, \tag{22.106}$$

where λ and q must be real since T^2 and T_3 are Hermitian operators. We start by noting that $T_\pm$ are step operators. We find:

$$T_3\,\{\,T_\pm\,|\,\lambda,q\,)\,\} = \{\,T_\pm\,T_3 + [\,T_3, T_\pm\,]\,\}\,|\,\lambda,q\,) = T_\pm\,\{\,T_3 + 1\,\}\,|\,\lambda,q\,) = (\,q+1\,)\,\{\,T_\pm\,|\,\lambda,q\,)\,\}\,, \tag{22.107}$$

so, assuming no degenerate eigenvectors,

$$T_+\,|\,\lambda,q\,) = A_+(\lambda,q)\,|\,\lambda,q+1\,)\,, \qquad \text{and} \qquad T_-\,|\,\lambda,q\,) = A_-(\lambda,q)\,|\,\lambda,q-1\,)\,. \tag{22.108}$$

Now we also have:

$$\langle\,\lambda,q\,|\,\frac{1}{2}\,\{\,T_+T_- + T_-T_+\,\}\,|\,\lambda,q\,) = \langle\,\lambda,q\,|\,\{\,T_3^2 - T^2\,\}\,|\,\lambda,q\,) = q^2 - \lambda \geq 0\,, \tag{22.109}$$

since the left-hand-side is positive definite.

Let us first consider the case when $\lambda \geq 0$. Then let $q_{\min} \geq 0$ be the minimum and $-q_{\max} \leq 0$ the maximum value of q, so that $q_{\min} \geq \sqrt{\lambda}$ and $q_{\max} \geq \sqrt{\lambda}$. Then, from Eqs. (22.108), q can have the values: $q = q_{\min}, q_{\min}+1, q_{\min}+2, \ldots$ and $q = -q_{\max}, -q_{\max}-1, -q_{\max}-2, \ldots$.

We next prove that $q_{\min} = q_{\max}$. The operator T_-, when acting on the lowest q-state $|\,\lambda, q_{\min}\,)$ must give zero:

$$T_-\,|\,\lambda,q_{\min}\,) = 0\,. \tag{22.110}$$

Operating again by T_+ gives:

$$T_+T_-\,|\,\lambda,q_{\min}\,) = \{\,T_3^2 - T_3 - T^2\,\}\,|\,\lambda,q_{\min}\,) = \{\,q_{\min}(q_{\min}-1) - \lambda\,\}\,|\,\lambda,q_{\min}\,) = 0\,, \tag{22.111}$$

from which we conclude that

$$\lambda = q_{\min}(q_{\min}-1) \geq 0\,. \tag{22.112}$$

Similarly, T_+ when acting on the highest state $|\,\lambda, -k_0\,)$ must give zero:

$$T_+\,|\,\lambda, -q_{\max}\,) = 0\,. \tag{22.113}$$

Operating again by T_- gives:

$$T_-T_+\,|\,\lambda, -q_{\max}\,) = \left\{\,T_3^2 + T_3 - T^2\,\right\}|\,\lambda, -q_{\max}\,) = \left\{\,q_{\max}(q_{\max}-1) - \lambda\,\right\}|\,\lambda, -q_{\max}\,) = 0\,, \tag{22.114}$$

from which we conclude that

$$\lambda = q_{\max}(q_{\max}-1) \geq 0\,. \tag{22.115}$$

Comparing Eqs. (22.115) and (22.115) we conclude that $q_{\min} = q_{\max} \equiv q_0 \geq 1$. So let us put $q_0 = k+1$, with $k \geq 0$, so that $\lambda = k(k+1)$. So now q can have the values: $q = k+1, k+2, k+3, \cdots$ and $q = -k-1, -k-2, -k-3, \cdots$. *The algebra places no restrictions on the value of* k, in particular k does not have to be integer or half-integer. This means that k is *not* an eigenvalue and must be regarded as a parameter of the eigenvector, which we now write as $|\,k, q\,)$.

To find $A_+(k,q)$, we take the inner product of the first of Eq. (22.108) with itself. This gives:

$$\begin{aligned} |\,A_+(k,q)\,|^2 &= (\,k,q\,|\,T_-T_+\,|\,k,q\,) = (\,k,q\,|\,\left\{\,T_3^2 + T_3 - T^2\,\right\}|\,k,q\,) \\ &= q(q+1) - k(k+1) = (q-k)\,(q+k+1)\,, \end{aligned} \tag{22.116}$$

So choosing the phase to be zero, we find:

$$A_+(k,q) = \sqrt{(q-k)\,(q+k+1)}\,. \tag{22.117}$$

Similarly, for $A_-(k,q)$, we have

$$\begin{aligned} |\,A_-(k,q)\,|^2 &= (\,k,q\,|\,T_+T_-\,|\,k,q\,) = (\,k,q\,|\,\left\{\,T_3^2 - T_3 - T^2\,\right\}|\,k,q\,) \\ &= q(q-1) - k(k+1) = (q+k)\,(q-k-1)\,. \end{aligned} \tag{22.118}$$

Again choosing the phase to be zero, we find:

$$A_-(k,q) = \sqrt{(q+k)\,(q-k-1)}\,. \tag{22.119}$$

For $\lambda < 0$, Eq. (22.109) only means that $q^2 \geq 0$, since q must be real. So there is no minimum or maximum for q. Eqs. (22.108) in this case means that q can have the values: $q_0, q_0 \pm 1, q_0 \pm 2, \ldots$, where now q_0 is arbitrary and unrelated to λ. The eigenvectors are orthonormal with respect to the q quantum number only:

$$(\,k,q\,|\,k,q'\,) = \delta_{q,q'}\,. \tag{22.120}$$

In particular, in general $(\,k,q\,|\,k',q\,) \neq 0$ if $k \neq k'$. This completes the proof. □

For the hydrogen Hamiltonian, we define three operators by:

$$T_1 = \frac{1}{2}\left\{\,R\,P_r^2 + \frac{\ell(\ell+1)}{R} - R\,\right\} \mapsto \frac{r}{2}\left\{\,-\frac{\partial^2}{\partial r^2} - \frac{2}{r}\,\frac{\partial}{\partial r} + \frac{\ell(\ell+1)}{r^2} - 1\,\right\}, \tag{22.121}$$

$$T_2 = R\,P_r \qquad\qquad \mapsto \frac{1}{i}\left\{\,r\,\frac{\partial}{\partial r} + 1\,\right\}$$

$$T_3 = \frac{1}{2}\left\{\,R\,P_r^2 + \frac{\ell(\ell+1)}{R} + R\,\right\} \mapsto \frac{r}{2}\left\{\,-\frac{\partial^2}{\partial r^2} - \frac{2}{r}\,\frac{\partial}{\partial r} + \frac{\ell(\ell+1)}{r^2} + 1\,\right\}, \tag{22.122}$$

which obey commutation rules Eqs. (22.100) of the $SO(2,1)$ algebra. For example, we see that:

$$\begin{aligned} [\,T_1, T_2\,] &= \frac{1}{2}\left\{\,[\,R\,P_r^2, R\,P_r\,] + \ell(\ell+1)\,[\,\frac{1}{R}, R\,P_r\,] - [\,R, R\,P_r\,]\,\right\} \\ &= \frac{1}{2}\left\{\,-i\,R\,P_r^2 - i\,\frac{\ell(\ell+1)}{R} - i\,R\,\right\} = -i\,T_3\,, \end{aligned} \tag{22.123}$$

as required. We leave proof of the other two commutation relations to the next exercise.

Exercise 78. Prove that Eqs. (22.121) satisfy commutation rules of the $SO(2,1)$ algebra, defined by Eqs. (22.100).

All the T_i operators are Hermitian with respect to a measure $\mu_1 = R$ (*not* R^2 !), in the sense that:

$$\begin{aligned}(\,\alpha\,|\,T_i\,\beta\,) &= \int_0^\infty r\,\mathrm{d}r\tilde{R}^*_\alpha(r)\,\big[\,T_i\,\tilde{R}_\beta(r)\,\big] \\ &= \int_0^\infty r\,\mathrm{d}r\big[\,T_i\,\tilde{R}_\alpha(r)\,\big]^*\,\tilde{R}_\beta(r) = (\,T_i\,\alpha\,|\,\beta\,)\,.\end{aligned} \tag{22.124}$$

Here we have denoted states for which the inner product is defined by Eq. (22.124) with measure μ_1 by $|\,\alpha\,)$ with parenthesis ends, and wave functions by a tilde: $\tilde{R}_\alpha(r) = (\,r\,|\,\alpha\,)$. We can see directly from the operator form of the T_i that they satisfy the relation: $T_i = \mu_1\,T^\dagger\,\mu_1^{-1}$.

From (22.121), we find that:

$$\begin{aligned}T_3 + T_1 &= R\,P_r^2 + \frac{\ell(\ell+1)}{R}\,, \\ T_3 - T_1 &= R\,.\end{aligned} \tag{22.125}$$

For our case, T^2 is given by:

$$\begin{aligned}T^2 = T_3^2 - T_1^2 - T_2^2 &= (T_3 - T_1)(T_3 + T_1) - [\,T_3, T_1\,] - T_2^2 \\ &= R^2\,P_r^2 + \ell(\ell+1) - i\,R\,P_r - R\,P_r\,R\,P_r = \ell(\ell+1)\,.\end{aligned} \tag{22.126}$$

So from Theorem 57, we write the eigenvalues and eigenvectors of T^2 and T_3 as:

$$\begin{aligned}T^2\,|\,\ell,q\,) &= \ell(\ell+1)\,|\,\ell,q\,)\,, \\ T_3\,|\,\ell,q\,) &= q\,|\,\ell,q\,)\,,\end{aligned} \tag{22.127}$$

with $q = \ell+1, \ell+2, \ell+3, \ldots$ with $\ell \geq 0$ and $q = -\ell-1, -\ell-2, -\ell-3\ldots$. Here we have written the eigenvectors as $|\,\ell,q\,)$ to indicated that they are orthogonal with respect to the measure μ. We shall see that not all of these eigenvectors are allowed for the physical Hamiltonian.

The operator $T_2 = RP_r$ generates a change of scale of the radial coordinate operator R and radial momentum operator P_r. That is since $i\,[\,T_2, R\,] = +1$ and $i\,[\,T_2, P_r\,] = -1$, the finite transformation $U(a)$ that does this is given by:[2]

$$\begin{aligned}U(a) = e^{iaT_2} \quad &\mapsto \exp\Big\{\,a\left(r\,\frac{\partial}{\partial r} + 1\right)\Big\}\,, \\ U^{-1}(a) = e^{-iaT_2} &\mapsto \exp\Big\{-a\left(r\,\frac{\partial}{\partial r} + 1\right)\Big\}\,.\end{aligned} \tag{22.128}$$

Then it is easy to show that:

$$U(a)\,R\,U^{-1}(a) = e^a\,R\,, \qquad U(a)\,P_r\,U^{-1}(a) = e^{-a}\,P_r\,. \tag{22.129}$$

The scale transformation preserves the commutation properties of R and P_r:

$$U(a)\,[\,R, P_r\,]\,U^{-1}(a) = [\,U(a)\,R\,U^{-1}(a), U(a)\,P_r\,U^{-1}(a)\,] = [\,R, P_r\,] = i\,, \tag{22.130}$$

and is unitary with respect to the measure μ_1. We can easily show this is a coordinate representation. We first note that

$$\begin{aligned}(\,r\,|\,U(a)\,\alpha\,) &= \exp\Big\{\,a\left(r\,\frac{\partial}{\partial r} + 1\right)\Big\}\,\tilde{R}_\alpha(r) = e^a\,\tilde{R}_\alpha(e^a\,r)\,, \\ (\,r\,|\,U^{-1}(a)\,\alpha\,) &= \exp\Big\{-a\left(r\,\frac{\partial}{\partial r} + 1\right)\Big\}\,\tilde{R}_\alpha(r) = e^{-a}\,\tilde{R}_\alpha(e^{-a}\,r)\,.\end{aligned} \tag{22.131}$$

[2]Sometimes this kind of transformation is called a **dilation** and the subsequent transformation of R a conformal transformation since the coordinate operator is expanded by a factor e^a.

On the other hand, $U^\dagger(a)$ is defined by:

$$\begin{aligned}(\,U^\dagger(a)\,\alpha\,|\,\beta\,) = (\,\alpha\,|\,U(a)\,\beta\,) &= \int_0^\infty r\,\mathrm{d}r\,\tilde{R}^*_\alpha(r)\,\exp\Big\{\,a\left(\,r\,\frac{\partial}{\partial r}+1\right)\Big\}\,\tilde{R}_\beta(r)\\ &= \int_0^\infty r\,\mathrm{d}r\,\tilde{R}^*_\alpha(r)\,e^a\,\tilde{R}_\alpha(e^a\,r) = \int_0^\infty r'\,\mathrm{d}r'\,\left[\,e^{-a}\,\tilde{R}^*_\alpha(e^{-a}\,r')\,\right]^*\,\tilde{R}_\beta(r')\\ &= \int_0^\infty r'\,\mathrm{d}r'\,\left[\,\Big\{\,-a\left(\,r'\,\frac{\partial}{\partial r'}+1\right)\Big\}\,\tilde{R}_\alpha(r')\right]^*\,\tilde{R}_\beta(r')\,,\end{aligned} \tag{22.132}$$

where we have changed variables: $r' = e^a r$. So $U(a)$ is unitary for the measure $\mu = r$. We can, of course, see this directly from the relation: $U^{-1}(a) = \mu_1\,U^\dagger(a)\,\mu_1^{-1}$.

Exercise 79. Prove Eqs. (22.129).

Now from Eq. (22.125), T_3+T_1 and T_3-T_1 transform under the $U(a)$ scale transformation by:

$$U(a)\,(\,T_3+T_1\,)\,U^{-1}(a) = e^{-a}\,(\,T_3+T_1\,)\,,\qquad U(a)\,(\,T_3-T_1\,)\,U^{-1}(a) = e^{a}\,(\,T_3-T_1\,)\,. \tag{22.133}$$

Of course T_2 is unchanged by the transformation. So T_1, T_2, and T_3 transform according to the rule:

$$U(a)\begin{pmatrix}T_1\\T_2\\T_3\end{pmatrix}U^{-1}(a) = \begin{pmatrix}\cosh a & 0 & -\sinh a\\ 0 & 1 & 0\\ -\sinh a & 0 & \cosh a\end{pmatrix}\begin{pmatrix}T_1\\T_2\\T_3\end{pmatrix}\,, \tag{22.134}$$

which resembles a rotation about the 2-axis by an imaginary angle ia.

First let us note that RH is Hermitian with respect to the measure μ_1 and that we can write RH in terms of the operators T_1 and T_3:

$$R\,H_\ell = \frac{R\,P_r^2}{2}+\frac{\ell(\ell+1)}{2R}-1 = \frac{1}{2}\,\Big\{\,T_3+T_1\,\Big\}-1\,. \tag{22.135}$$

We write the eigenvalue equation for the hydrogen Hamiltonian H as:

$$H\,|\,n,\ell\,\rangle = \epsilon_n\,|\,n,\ell\,\rangle\,, \tag{22.136}$$

which is Hermitian with respect to measure μ_2. Here $|\,n,\ell\,\rangle$, written in using angles, are written in coordinate space as $R_{n,\ell}(r) = \langle\,r\,|\,n,\ell\,\rangle$ and normalized with the measure μ_2. That is:

$$\langle\,n,\ell\,|\,n',\ell'\,\rangle = \int_0^\infty R^*_{n,\ell}(r)\,R_{n',\ell'}(r)\,r^2\,\mathrm{d}r\,. \tag{22.137}$$

The eigenvalue problem for the measure μ_1 is given by multiplying Eq. (22.136) by R. This gives:

$$R\,(\,H_\ell-\epsilon_n\,)\,|\,n,\ell\,\rangle = 0\,,\qquad \epsilon_n = -\frac{1}{2\,n^2}\,. \tag{22.138}$$

From Eqs. (22.125) and (22.135), we find:

$$\begin{aligned}R\,(\,H-\epsilon_n\,) &= \frac{R\,P_r^2}{2}+\frac{\ell(\ell+1)}{2R}-1+\frac{R}{2\,n^2}\\ &= \frac{1}{2}\,\Big\{\,T_3+T_1\,\Big\}+\frac{1}{2\,n^2}\,\Big\{\,T_3-T_1\,\Big\}-1\\ &= \frac{1}{2}\Big\{\,1+\frac{1}{n^2}\,\Big\}\,T_3+\frac{1}{2}\Big\{\,1-\frac{1}{n^2}\,\Big\}\,T_1-1\,.\end{aligned} \tag{22.139}$$

The trick is now to use the scale transformation (22.134) about the 2-axis to "rotate" the operator $R\,(H-\epsilon_n)$ to "point" in the 3-direction. This gives:

$$\begin{aligned} U(a)\,R\,(\,H-\epsilon_n\,)\,U^{-1}(a) &= \frac{e^{-a}}{2}\,\{\,T_3+T_1\,\}+\frac{e^{a}}{2\,n^2}\,\{\,T_3-T_1\,\}-1 \\ &= \frac{1}{2}\left\{\,e^{-a}+\frac{e^{a}}{n^2}\,\right\}T_3+\frac{1}{2}\left\{\,e^{-a}-\frac{e^{a}}{n^2}\,\right\}T_1-1\,. \end{aligned} \tag{22.140}$$

T_1 can now be eliminated from the rotated Hamiltonian operator by choosing a such that:

$$e^a=n\,,\qquad a=\ln(n)=-\frac{1}{2}\,\ln\!\Big(\frac{1}{n^2}\Big)=-\frac{1}{2}\,\ln(-2\epsilon_n)\,. \tag{22.141}$$

Then Eq. (22.140) becomes:

$$U(a)\,R\,(\,H-\epsilon_n\,)\,U^{-1}(a)=e^{-a}\,T_3-1=\frac{1}{n}\,(\,T_3-n\,)\,. \tag{22.142}$$

Then we find:

$$\frac{1}{n}\,(\,T_3-n\,)\,\{\,U(a)\,|\,n,\ell\,\rangle\,\}=U(a)\,R\,(\,H-\epsilon_n\,)\,|\,n,\ell\,\rangle=0\,. \tag{22.143}$$

So $U(a)\,|\,n,\ell\,\rangle$ is a common eigenvector of T_3 with eigenvalue n, and T with eigenvalue $\ell(\ell+1)$. Comparing (22.143) with the eigenvalue equations for T^2 and T_3, Eqs. (22.127), we see that

$$|\,n,\ell\,\rangle=\mathcal{N}_n\,U^{-1}(a)\,|\,\ell,n\,)\,, \tag{22.144}$$

where $\mathcal{N}_n$ is a normalization factor, and where $k=\ell$ and $q=n=\ell+1,\ell+2,\ell+3,\ldots>0$. That is, the positive eigenvalues of T_3 are the principle quantum numbers of the hydrogen Hamiltonian. The negative eigenvalues of T_3 are not physically realized here, which means that the states $|\,\ell,n\,)$ *are not complete!*

The duel of Eq. (22.144) is:

$$\langle\,n,\ell\,|=\mathcal{N}_n^*\,(\,\ell,n\,|\,U(a)\,R\,, \tag{22.145}$$

the extra factor of R coming from the difference between the measures μ_1 and μ_2. So the normalization is fixed by the requirement:

$$\begin{aligned} \langle\,n,\ell\,|\,n,\ell\,\rangle &= |\,\mathcal{N}_n\,|^2\,(\,\ell,n\,|\,U(a)\,R\,U^{-1}(a)\,|\,\ell,n\,) \\ &= |\,\mathcal{N}_n\,|^2\,e^a\,(\,\ell,n\,|\,R\,|\,\ell,n\,) \\ &= |\,\mathcal{N}_n\,|^2\,e^a\,(\,\ell,n\,|\,\{\,T_3-(T_+ + T_-)/2\,\}\,|\,\ell,n\,)=|\,\mathcal{N}_n\,|^2\,n^2=1\,. \end{aligned} \tag{22.146}$$

So we find $\mathcal{N}_n=1/n$. This does *not* mean that $\langle\,n,\ell\,|\,n,\ell'\,\rangle=0$, and in fact it is not zero. The radial wave functions in coordinate space in the two basis sets $R_{n,\ell}(r)=\langle\,r\,|\,n,\ell\,\rangle$ and $\tilde{R}_{\ell,n}(r)=\langle r|\ell,n)$ are related by:

$$\begin{aligned} R_{n,\ell}(r)=\langle\,r\,|\,n,\ell\,\rangle &= \mathcal{N}_n\,\langle\,r\,|\,e^{-iaRP_r}\,|\,\ell,n\,)=\mathcal{N}_n\,\exp\!\Big\{-a\Big(\,r\frac{\partial}{\partial r}+1\Big)\,\Big\}\,\tilde{R}_{\ell,n}(r) \\ &= \mathcal{N}_n\,e^{-a}\,\tilde{R}_{\ell,n}(\,e^{-a}r\,)=\frac{1}{n^2}\,\tilde{R}_{\ell,n}(\,r/n\,)\,. \end{aligned} \tag{22.147}$$

Note that $|\tilde{R}_{\ell,n}(r)|^2\,r\,\mathrm{d}r$ is *not* the probability of finding the electron between r and $r+\mathrm{d}r$, because of the scaling factor r/n.

Radial wave functions $\tilde{R}_{\ell,n}(r)$ in the $|\,\ell,n\,)$ basis are readily obtained by operation by T_- on the lowest state, and then using T_+ to obtain the rest of the wave functions. The ground state $|\,\ell,\ell+1\,)$ basis is given by the solution of:

$$\begin{aligned} T_-\,|\,\ell,\ell+1\,) &= \{\,T_1-i\,T_2\,\}\,|\,\ell,\ell+1\,)=-\{\,T_3-T_1+i\,T_2-T_3\,\}\,|\,\ell,\ell+1\,) \\ &= -\{\,i\,R\,P_r+R-\ell\,\}\,|\,\ell,\ell+1\,)=0\,. \end{aligned} \tag{22.148}$$

In coordinate space, this reads:

$$\left\{ r\,\frac{\partial}{\partial r} + r - \ell \right\} \tilde{R}_{\ell,\ell+1}(r) = 0\,, \tag{22.149}$$

which has the solution,

$$\tilde{R}_{\ell,\ell+1}(r) = \frac{2^{\ell+1}}{\sqrt{(2\ell+1)!}}\, r^{\ell}\, e^{-r}\,, \tag{22.150}$$

which has been normalized to one with measure μ_1. From (22.147), the radial wave function $R_{n,n-1}(r)$ for $\ell = n-1$ is then given by:

$$R_{n,n-1}(r) = \frac{1}{n^2}\,\tilde{R}_{n-1,n}(r/n) = \sqrt{\frac{2^{2n+1}}{n^{2n+1}\,(2n)!}}\, r^{n-1}\, e^{-r/n}\,, \tag{22.151}$$

in agreement with our result Eq. (22.92) which we found in Section 22.3.4 using the operator factorization method. Radial wave functions for arbitrary values of ℓ and n can be found by application of the n-raising operator T_+ on the state $|\,\ell,\ell+1\,\rangle$.

Exercise 80. Show that the radial wave functions $\tilde{R}_{\ell,n}(r)$, normalized with the measure μ_1, are given in general by:

$$\tilde{R}_{\ell,n}(r) = \frac{2^{\ell+1}}{\sqrt{(n+\ell)!\,(n-\ell-1)!}}\,\frac{e^r}{r^{\ell+1}}\,\frac{\partial^{n-\ell-1}}{\partial r^{n-\ell-1}}\left[\, r^{n+\ell}\, e^{-2r}\,\right]. \tag{22.152}$$

From Eqs. (22.144) and (22.145), the matrix elements of an operator O can be computed in either basis by means of the relation:

$$\langle\, n,\ell\,|\,O\,|\,n',\ell'\,\rangle = \frac{1}{n^2}\,(\,\ell,n\,|\,U(a)\,RO\,U^{-1}(a)\,|\,\ell',n'\,) = (\,\ell,n\,|\,\widetilde{RO}\,|\,\ell',n'\,)\,, \tag{22.153}$$

where

$$\widetilde{RO} = \frac{1}{n^2}\,U(a)\,RO\,U^{-1}(a)\,. \tag{22.154}$$

22.3.6 $SO(4,2)$ algebra

In Section 22.3.3, we developed

22.3.7 The fine structure of hydrogen

We have found that the Hamiltonian for the hydrogen atom, which we now call H_0, in ordinary units is given by:

$$H_0 = \frac{P^2}{2m} - \frac{e^2}{R}\,, \tag{22.155}$$

with energies and eigenvectors given by:

$$E_{0\,n} = \frac{1}{2}\,mc^2\alpha^2\,\frac{1}{n^2}\,, \qquad \text{and} \qquad |\,n,\ell,m_\ell\,\rangle\,. \tag{22.156}$$

A careful examination of the spectra of hydrogen, however, reveals that Eq. (22.156) is only approximately correct, but that there is a *fine structure* to the energy spectra which is not accounted for by this equation. This turns out to be due to the relativistic nature of the electron. One of these relativistic manifestations is that the electron has an intrinsic spin, with the value $s = 1/2$, and an intrinsic magnetic moment. The other manifestation is that the kinetic energy of the electron needs to be corrected for the relativistic mass change with velocity. We discuss both of these corrections in this section.

Spin-orbit force

The spin-orbit force for the electron in hydrogen is a relativistic interaction between the magnetic moment of the electron and the effective magnetic field as seen by the electron in its rest frame as a result of the electric field of the proton. Sometimes this effect is described by noting that in the rest frame of the electron, the proton is orbiting the electron and so creates a magnetic field at the position of the electron; however, there is a subtle correction to this simple explanation due to a relativistic procession of the electron, called the Thomas precession. So the spin-orbit energy is given by:

$$H_{\rm so} = -\boldsymbol{\mu}_{\rm e} \cdot \mathbf{B}_{\rm eff} \,, \qquad \boldsymbol{\mu}_{\rm e} = -\frac{e}{mc}\,\mathbf{S} \,, \qquad \mathbf{B}_{\rm eff} = -\frac{1}{2c}\,\mathbf{V} \times \mathbf{E} \,, \tag{22.157}$$

where $\mathbf{S}$ is the spin of the electron, and $\mathbf{E}$ is the electric field due to the proton, given by:

$$\mathbf{E} = \frac{e\,\mathbf{R}}{R^3} \,. \tag{22.158}$$

Putting this together, and noting that $\mathbf{R} \times \mathbf{V} = \mathbf{L}/m$, we find:

$$H_{\rm so} = \frac{e^2}{2m^2c^2}\,\frac{\mathbf{L}\cdot\mathbf{S}}{R^3} = \frac{1}{2}\,mc^2\,\alpha^4 \left(\frac{a}{R}\right)^3 \frac{\mathbf{L}\cdot\mathbf{S}}{\hbar^2} \,. \tag{22.159}$$

Writing this in atomic units with lengths in units of a, angular momentum in units of $\hbar$, and energies in units of $E_0 = mc^2\alpha^2$, the spin-orbit Hamiltonian is:

$$\bar{H}_{\rm so} = \frac{\alpha^2}{2}\,\frac{\mathbf{L}\cdot\mathbf{S}}{\bar{R}^3} \,, \tag{22.160}$$

so this energy is down from the energies of the major shells by a factor of α^2. For an derivation of the effective magnetic field $\mathbf{B}_{\rm eff}$, including Thomas presession, see the book by J. D. Jackson. With the extra spin degree of freedom, we can construct direct product states given by $|\, n, \ell, m_\ell, s, m_s \,\rangle$, or form the coupled states:

$$|\, n, (\ell, s)\, j, m_j \,\rangle = \sum_{m_\ell, m_s} \langle\, \ell, m_\ell, s, m_s \,|\, (\ell, s)\, j, m_j \,\rangle\, |\, n, \ell, m_\ell, s, m_s \,\rangle \,. \tag{22.161}$$

Both of these states are eigenstates of H_0. However, the coupled state, defined by (22.161), is also an eigenstate of the spin-orbit force. We can easily calculate the spin-orbit energies using these states. Since $\mathbf{J} = \mathbf{L} + \mathbf{S}$, we have:

$$\mathbf{L}\cdot\mathbf{S} = \frac{1}{2}\,(\,J^2 - L^2 - S^2\,) \,, \tag{22.162}$$

so that we find:

$$\langle\, (\ell, s)\, j, m_j \,|\, \mathbf{L}\cdot\mathbf{S} \,|\, (\ell, s)\, j, m_j \,\rangle = \frac{1}{2}\big(\, j(j+1) - \ell(\ell+1) - 3/4 \,\big) \,. \tag{22.163}$$

We show an alternate way to calculate this matrix element in Section 21.6.1 using fancy angular momentum technology (which is *not* necessary in this case!). Expectation values of the radial function in atomic units is given by:

$$\left\langle \frac{1}{\bar{R}^3} \right\rangle_{n,\ell} = \frac{2}{n^3\,\ell(\ell+1)(2\ell+1)} \,. \tag{22.164}$$

So using first order perturbation theory, the energy shift in atomic units in hydrogen due to the spin-orbit force is given by:

$$\begin{aligned} \Delta\bar{E}_{({\rm so})\,n,\ell,j} &= \langle\, n, (\ell, s)\, j, m_j \,|\, \bar{H}_{\rm so} \,|\, n, (\ell, s)\, j, m_j \,\rangle = \frac{\alpha^2}{2}\,\frac{j(j+1) - \ell(\ell+1) - 3/4}{n^3\,\ell(\ell+1)(2\ell+1)} \\ &= (\pm)\,\alpha^2\,\frac{1}{n^3\,(2\ell+1)(2j+1)} \,, \end{aligned} \tag{22.165}$$

with the (+) for $j = \ell + 1/2$ and (−) for $j = \ell - 1/2$. Note that the spin-orbit energy is *finite* for s-states even though $\mathbf{L}\cdot\mathbf{S}$ vanishes for s-states and $\langle\, 1/r^3 \,\rangle$ blows up. So we have cheated here, canceling the divergences, and obtaining something finite. The resolution of this strange result is to carefully examine the spin-orbit interaction for s-states, averaging the Hamiltonian over a small region about the origin. This gives a "contact" interaction proportional to a δ-function which should be added to the spin-orbit interaction. The end result of the analysis gives the same answer we found in Eq. (22.165). For details, see ref. XX.

Relativistic correction

The relativistic expansion of the total energy E of an electron with momentum p is given by:

$$H = \sqrt{P^2c^2 + (mc^2)^2} \approx mc^2 + \frac{P^2}{2\,m} - \frac{P^4}{8\,m^3c^2} + \cdots \tag{22.166}$$

We accounted for the second term in this expansion in our non-relativistic expression for the Hamiltonian for hydrogen, but we neglected the first and third terms. The first term is just a constant energy, which we will ignore. However the third term gives a relativistic correction to the energy, which we will call

$$H_{\rm rel} = -\frac{P^4}{8\,m^3c^2} = -\frac{\hbar^4}{8\,m^3c^2a^4}\left[\frac{a\,P}{\hbar}\right]^4 = -\frac{1}{8}\,mc^2\,\alpha^4\left[\frac{a\,P}{\hbar}\right]^4 . \tag{22.167}$$

So in atomic units, with $P = \hbar\,\bar{P}/a$ and energies in units of $E_0 = mc^2\alpha^2$, the relativistic correction Hamiltonian is given by:

$$\bar{H}_{\rm rel} = -\frac{\alpha^2}{8}\,\bar{P}^4\,, \tag{22.168}$$

which is the same order of magnitude as the spin-orbit force, down from the major shell energy by a factor of α^2. The complete fine structure Hamiltonian is then given by the sum:

$$\bar{H}_{\rm fs} = \bar{H}_{\rm so} + \bar{H}_{\rm rel}\,. \tag{22.169}$$

We can easily calculate the energy shift due to H_{rel} using the unperturbed Hamilton $\bar{H}$ for hydrogen given in Eq. (22.27):

$$\bar{P}^2 = 2\,\bar{H} + \frac{2}{\bar{R}}\,, \qquad \text{with} \qquad \bar{H}\,|\,n,(\ell,s)\,j,m_j\,\rangle = -\frac{1}{2\,n^2}\,|\,n,(\ell,s)\,j,m_j\,\rangle\,. \tag{22.170}$$

Substituting this into (22.168) gives:

$$\bar{H}_{\rm rel} = -\frac{\alpha^2}{2}\left\{\,\bar{H}^2 + \bar{H}\,\frac{1}{\bar{R}} + \frac{1}{\bar{R}}\,\bar{H} + \frac{1}{\bar{R}^2}\,\right\}, \tag{22.171}$$

and expectation values of this in the states $|\,n,(\ell,s)\,j,m_j\,\rangle$ gives:

$$\begin{aligned}\Delta\bar{E}_{({\rm rel})\,n,\ell,j} &= \langle\, n,(\ell,s)\,j,m_j\,|\,\bar{H}_{\rm rel}\,|\,n,(\ell,s)\,j,m_j\,\rangle \\ &= -\frac{\alpha^2}{8}\left\{\,\frac{1}{n^4} - \frac{4}{n^2}\left\langle\frac{1}{\bar{R}}\right\rangle_{n,\ell} + 4\left\langle\frac{1}{\bar{R}^2}\right\rangle_{n,\ell}\right\}\end{aligned} \tag{22.172}$$

Now

$$\left\langle\frac{1}{\bar{R}}\right\rangle_{n,\ell} = \frac{1}{n^2}\,, \qquad \text{and} \qquad \left\langle\frac{1}{\bar{R}^2}\right\rangle_{n,\ell} = \frac{2}{n^3\,(2\ell+1)}\,. \tag{22.173}$$

Substitution into (22.172) gives:

$$\Delta\bar{E}_{({\rm rel})\,n,\ell,j} = \frac{\alpha^2}{2}\left\{\,\frac{3}{4}\,\frac{1}{n^4} - \frac{2}{n^3\,(2\ell+1)}\,\right\}. \tag{22.174}$$

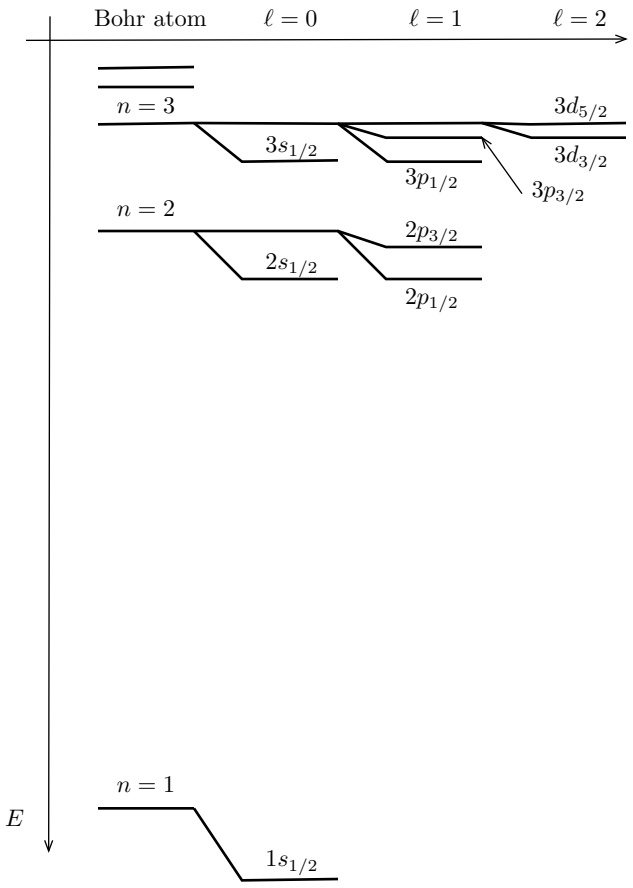

Figure 22.1: The fine structure of hydrogen (not to scale). Levels with the same value of j are degenerate.

Adding the results of Eqs. (22.174) and (22.165) gives the fine structure splitting energy:

$$\Delta\bar{E}_{\rm fs} = \Delta\bar{E}_{({\rm so})\,n,\ell,j} + \Delta\bar{E}_{({\rm rel})\,n,\ell,j} = -\frac{\alpha^2}{2\,n^4}\left(\frac{n}{j+1/2} - \frac{3}{4}\right), \tag{22.175}$$

which is now independent of ℓ. The total energy of hydrogen in ordinary units, including the mass-energy and relativistic effects to first order, is now given by:

$$E_{n,j} = mc^2\left\{1 - \frac{\alpha^2}{2n^2} - \frac{\alpha^4}{2n^4}\left(\frac{n}{j+1/2} - \frac{3}{4}\right) + \cdots\right\}, \tag{22.176}$$

which is in agreement with the expansion in powers of α of the exact energy from the solution of the relativistic Dirac equation for hydrogen (see, for example, Bjorken and Drell [11]), and also with experiment. Notice that the energy levels to this order depend only on n and j and are independent of ℓ, so that states with the same value of j but different values of ℓ, such as $2s_{1/2}$ and $2p_{1/2}$ states, have the same energy. This odd degeneracy, which is somehow due to the $O(4)$ symmetry, persists even for the exact relativistic Dirac equation. The energy levels, including the fine structure, are shown in Fig. 22.1. Note all the states are shifted *lower* by the hyperfine energy and that the $j = \ell + 1/2$ state is *above* the $j = \ell - 1/2$ state.

The energy difference between the $n = 2$ and $n = 1$ major shells in hydrogen is $(3/4)\,E_R = 10.21$ eV corresponding to a frequency of $(3/4)\,f_R = 2.468 \times 10^9$ MHz or a wavelength of $\lambda = (4/3)\,\lambda_R = 121$ nm, whereas the energy difference between the $2p_{3/2}$ and $2p_{1/2}$ fine structure levels is $(\alpha^2/16)\,E_R = 4.530 \times 10^{-5}$ eV or a frequency of $(\alpha^2/16)\,f_R = 10,950$ MHz.

We next see if the ℓ-degeneracy persists after taking into account interactions between the magnetic moment of the proton and the electron.

22.3.8 The hyperfine structure of hydrogen

The hyperfine structure of the atomic energy levels of hydrogen is due to the interaction of the magnetic moment of the *proton* with magnetic fields due to the electron. It comes in two parts: (1) the dipole-dipole interaction between the magnetic dipoles of the proton and the electron, and (2) the interaction between the magnetic dipole of the the proton and a magnetic field created by the orbital motion of the electron in the atom. We find the interaction energy of these effects in order.

For the dipole-dipole interaction, we note first that the magnetic moments of the electron and proton is given by:

$$\boldsymbol{\mu}_{\rm e} = -\frac{e}{mc}\,\mathbf{S}_{\rm e}\,, \qquad \boldsymbol{\mu}_{\rm p} = +\lambda_{\rm p}\,\frac{e}{Mc}\,\mathbf{S}_{\rm p}\,, \tag{22.177}$$

where e is the *magnitude* of the charge on of the electron, m is the mass of the electron, M the mass of the proton, and $\mathbf{S}_{\rm e}$ and $\mathbf{S}_{\rm p}$ the spin of the electron and proton respectively. The anomalous magnetic moment of the proton is found experimentally to give $\lambda_{\rm p} = +2.793$. So the dipole-dipole energy is:

$$\begin{aligned} H_{dd} &= \frac{1}{R^3}\left\{\,\boldsymbol{\mu}_{\rm e}\cdot\boldsymbol{\mu}_{\rm p} - 3\,(\,\boldsymbol{\mu}_{\rm e}\cdot\hat{\mathbf{R}}\,)\,(\,\boldsymbol{\mu}_{\rm p}\cdot\hat{\mathbf{R}}\,)\,\right\}, \\ &= -\frac{\lambda_{\rm p} e^2\,\hbar^2}{mM\,c^2\,a^3}\left(\frac{a}{R}\right)^3\left\{\,\mathbf{S}_{\rm e}\cdot\mathbf{S}_{\rm p} - 3\,(\,\mathbf{S}_{\rm e}\cdot\hat{\mathbf{R}}\,)\,(\,\mathbf{S}_{\rm p}\cdot\hat{\mathbf{R}}\,)\,\right\}/\hbar^2\,, \\ &= -mc^2\,\alpha^4\,\lambda_{\rm p}\left(\frac{m}{M}\right)\frac{1}{\bar{R}^3}\left\{\,\mathbf{S}_{\rm e}\cdot\mathbf{S}_{\rm p} - 3\,(\,\mathbf{S}_{\rm e}\cdot\hat{\mathbf{R}}\,)\,(\,\mathbf{S}_{\rm p}\cdot\hat{\mathbf{R}}\,)\,\right\}/\hbar^2\,, \end{aligned} \tag{22.178}$$

where $\hat{\mathbf{r}}$ is a unit vector from the proton to the electron and $\bar{r} = r/a$ is the coordinate of the electron in atomic units in the center of mass system. So in atomic units, the dipole-dipole Hamiltonian is:

$$\bar{H}_{dd} = \alpha^2\,\lambda_{\rm p}\left(\frac{m}{M}\right)\frac{1}{\bar{R}^3}\left\{\,\mathbf{S}_{\rm e}\cdot\mathbf{S}_{\rm p} - 3\,(\,\mathbf{S}_{\rm e}\cdot\hat{\mathbf{R}}\,)\,(\,\mathbf{S}_{\rm p}\cdot\hat{\mathbf{R}}\,)\,\right\}. \tag{22.179}$$

So this energy is down by a factor of $2\lambda_{\rm p}\,(m/M) \approx 1/329$ from the fine structure splitting energy.

For the dipole-orbit energy, using the Biot-Savart law, the magnetic field $\mathbf{B}$ at the proton as a result of electron motion about the proton (at the origin in the center of mass system) is given by:

$$\mathbf{B} = \frac{(-e\,\mathbf{V})\times(-\mathbf{R})}{R^3} = \frac{e}{m}\,\frac{\mathbf{P}\times\mathbf{r}}{R^3} = -\frac{e}{m}\,\frac{\mathbf{L}}{R^3}\,, \tag{22.180}$$

where $\mathbf{L}$ is the angular momentum of the electron in the center of mass system. So the energy of this interaction of the magnetic field of the electron with the proton at the origin is given by:

$$H_{do} = -\boldsymbol{\mu}_{\rm p}\cdot\mathbf{B} = \frac{\lambda_{\rm p} e^2\,\hbar^2}{mM\,c^2\,a^3}\left(\frac{a}{R}\right)^3\mathbf{L}\cdot\mathbf{S}_{\rm p}/\hbar^2 = mc^2\,\alpha^4\,\lambda_{\rm p}\left(\frac{m}{M}\right)\frac{1}{\bar{R}^3}\,\mathbf{L}\cdot\mathbf{S}_{\rm p}/\hbar^2\,. \tag{22.181}$$

So measuring the energy in terms of the Bohr energy, the dipole-orbit Hamiltonian is:

$$\bar{H}_{do} = \alpha^2\,\lambda_{\rm p}\left(\frac{m}{M}\right)\frac{\mathbf{L}\cdot\mathbf{S}_{\rm p}}{\bar{R}^3}\,. \tag{22.182}$$

This is of the same order as the dipole-dipole term, so the final hyperfine splitting Hamiltonian in atomic units is given by:

$$\bar{H}_{\rm hf} = \alpha^2\,\lambda_{\rm p}\left(\frac{m}{M}\right)\frac{\mathbf{K}_{\rm e}\cdot\mathbf{S}_{\rm p}}{\bar{R}^3}\,, \qquad \mathbf{K}_{\rm e} = \mathbf{L} - \mathbf{S}_{\rm e} + 3\,(\mathbf{S}_{\rm e}\cdot\hat{\mathbf{R}})\,\hat{\mathbf{R}}\,. \tag{22.183}$$

Now that we are including the spin of the proton in the dynamics of the atom, there is an additional two degrees of freedom. So now the total angular momentum of the atom $\mathbf{F}$ is given by the sum:

$$\mathbf{F} = \mathbf{L} + \mathbf{S}_{\rm e} + \mathbf{S}_{\rm p} \,. \tag{22.184}$$

We will find that if we couple these angular momentum states in the following way:

$$|\, n, (\ell, s_e)\, j, s_p, f, m_f \,\rangle \,, \tag{22.185}$$

matrix elements of the hyperfine splitting Hamiltonian, given in Eq. (22.183), are diagonal and independent of m_f. We work out the details in Section 21.6.3. There, we show that the hyperfine energy shift is diagonal in f, m_f, and ℓ, independent of m_f, and in atomic units, is given by:

$$\Delta\bar{E}_{\ell,j,f} = \alpha^2\, \lambda_{\rm p} \left(\frac{m}{M}\right) \frac{f(f+1) - j(j+1) - 3/4}{n^3\, j(j+1)\,(2\ell+1)} \,. \tag{22.186}$$

where $f = j \pm 1/2$. The energy level diagram for the $n = 1$ and $n = 2$ levels of hydrogen, including the hyperfine energy, is shown in Fig. 22.2. The $n = 1$ state of hydrogen is now split into two parts with total angular momentum $f = 0$ and $f = 1$. The $f = 0$ ground state is shifted *lower* by a factor of 6/3 whereas the $f = 1$ state is shifted *higher* by a factor of 2/3, so the splitting of this state ΔE, in frequency units, is given:

$$f = \Delta E/(2\pi\hbar) = f_R\, 2\alpha^2\, \lambda_{\rm p} \left(\frac{m}{M}\right) \frac{8}{3} = 1421\,{\rm MHz} \,, \qquad \text{or} \qquad \lambda = c/f = 21\,{\rm cm}. \tag{22.187}$$

and is the 21 cm radiation seen in the absorption spectra from the sun. For the electron in the $1s_{1/2}$ state, the electron and the proton have spins pointing in *opposite* directions in the $f = 0$ ground state, whereas the $f = 1$ state has spins pointing in the same direction, giving rise to a three-fold degeneracy. The hyperfine splitting between the $2s_{1/2}$ levels is 178 MHz, between the $2p_{1/2}$ levels is 59.2 MHz, and between the $2p_{3/2}$ levels is 11.8 MHz. This leaves a small splitting between the $2s_{1/2}$ and the $2p_{1/2}$ $f = 0$ states of about 89 MHz. The measurement of the splitting between these states was a major experimental effort in the 1950's. A precise measurement by Lamb (ref?), found that there was a shift in energy of 1058 MHz between the $2s_{1/2}$ and $2p_{1/2}$ states which was not explainable in terms of ordinary quantum mechanics. The theoretical calculation of this Lamb shift was finally carried out using quantum field theory methods, and is still regarded as one of the major achievements of quantum field theory, a topic we do not discuss here.

22.3.9 The Zeeman effect

The Zeeman effect is a splitting of the energy levels of atoms as a result of the action of a static and position independent magnetic field $\mathbf{B}$ with the electrons.

For hydrogen, adding the fine structure Hamiltonian from Eq. (22.169) and the hyperfine Hamiltonian from Eq. (22.183), the complete Hamiltonian for hydrogen is now given by:

$$H = \frac{(\mathbf{P} + (e/c)\,\mathbf{A})^2}{2m} - \frac{e^2}{R} - \boldsymbol{\mu}_e \cdot \mathbf{B} + H_{\rm fs} + H_{\rm hf} \,, \qquad \boldsymbol{\mu}_e = -\frac{e}{mc}\,\mathbf{S} \,, \tag{22.188}$$

where e is again the *magnitude* of the charge on of the electron and $\boldsymbol{\mu}_e$ is the intrinsic magnetic moment of the electron with $\mathbf{S}$ the electron spin. For a constant magnetic field,

$$\mathbf{A} = \frac{1}{2}\,\mathbf{B} \times \mathbf{R} \,, \qquad \text{so that} \qquad \mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A} \,. \tag{22.189}$$

Now since $[\,\mathbf{P}, \mathbf{A}\,] = 0$ and $(\mathbf{P} \cdot \mathbf{A}) = 0$, the expansion of Eq. (22.188) gives:

$$\begin{aligned} H &= \frac{P^2}{2m} - \frac{e^2}{R} + \frac{e}{2\,m\,c}\,(\mathbf{B} \times \mathbf{R}) \cdot \mathbf{P} + \frac{e}{m\,c}\,\mathbf{S} \cdot \mathbf{B} + H_{\rm fs} + H_{\rm hf} + \cdots \\ &= H_0 + H_{\rm fs} + H_{\rm hf} + \frac{e}{2\,m\,c}\,(\,\mathbf{L} + 2\,\mathbf{S}\,) \cdot \mathbf{B} + \cdots \\ &= H_0 + H_{\rm fs} + H_{\rm hf} + H_{\rm z} + \cdots \end{aligned} \tag{22.190}$$

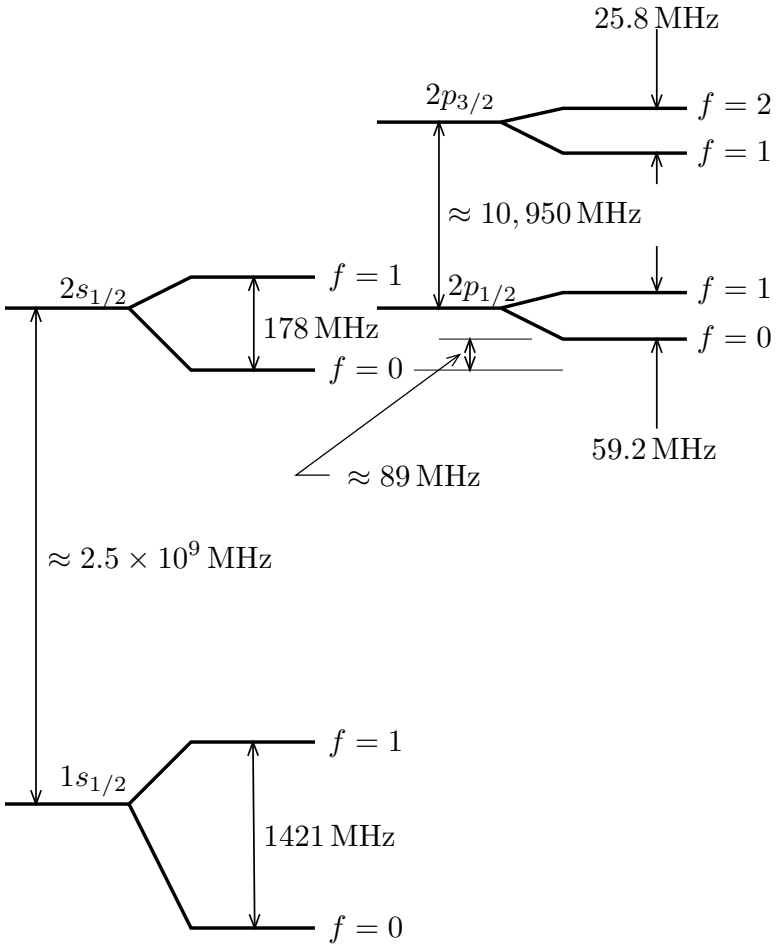

Figure 22.2: The hyperfine structure of the $n = 1$ and $n = 2$ levels of hydrogen (not to scale).

where the Zeeman Hamiltonian for hydrogen H_z is given by:

$$H_z = \mu_B \left(\mathbf{L} + 2\,\mathbf{S} \right) \cdot \mathbf{B}/\hbar \,, \qquad \text{with} \qquad \mu_B = \frac{e\,\hbar}{2\,mc} \,. \tag{22.191}$$

Here μ_B is called the Bohr magneton, and has a value of 1.40 MHz/Gauss (See Table A.1 in appendix A). We have ignored the quadratic magnetic field term in the Hamiltonian and the effect of the magnetic field on the magnetic moment of the proton, since it is smaller than that of the electron by a factor of $\lambda_p\,(m/M)$.

If we take the z-axis to be in the direction of the field $\mathbf{B}$, then the Zeeman Hamiltonian is diagonal in the states: $|\,\ell, m_\ell, s_e, m_{s_e}, s_p, m_{s_p}\,\rangle$. That is:

$$\langle\,\ell, m_\ell, s_e, m_{s_e}, s_p, m_{s_p}\,|\,H_z\,|\,\ell', m'_\ell, s_e, m'_{s_e}, s_p, m'_{s_p}\,\rangle = \delta_{\ell,\ell'}\delta_{m_\ell,m'_\ell}\delta_{m_{s_p},m'_{s_p}}\,\mu_B B\,(\,m_\ell + 2\,m_{s_e}\,)\,, \tag{22.192}$$

and proportional to the strength of the magnetic field B. If the magnetic field is so strong that Zeeman energy is much larger than the fine structure and hyperfine structure, then the fine and hyperfine structure are perturbations on top of the energy shifts given by Eq. (22.192). This takes a very large field. If, however, the Zeeman energy is less that or on the order of the hyperfine energy, we must diagonalize the sub-matrix for the hyperfine Hamiltonian along with the Zeeman Hamiltonian. For this diagonalization, it is easier to carry out the diagonalization using the coupled vectors: $|\,(\ell, s_e)\,j, s_p, f, m_f\,\rangle$. The factor of 2 multiplying the

spin term complicates the calculation of the energy shift for the Zeeman Hamiltonian using coupled states. We do this calculation in Section 21.6.4. For the case of the $1s_{1/2}$ hyperfine levels, using a simplified notation $|\, f, m_f \,\rangle$ for these states, we find the matrix elements:

$$\begin{aligned} \langle\, 0,0 \,|\, H_z \,|\, 0,0 \,\rangle = \langle\, 1,0 \,|\, H_z \,|\, 1,0 \,\rangle = \mu_B B \,, \\ \langle\, 1,1 \,|\, H_z \,|\, 1,1 \,\rangle = -\langle\, 1,1 \,|\, H_z \,|\, 1,1 \,\rangle = \langle\, 0,0 \,|\, H_z \,|\, 1,0 \,\rangle = \langle\, 1,0 \,|\, H_z \,|\, 0,0 \,\rangle = \mu_B B \,. \end{aligned} \tag{22.193}$$

So if we put E_0 to be the energy of the $f = 0$ hyperfine state and E_1 the $f = 1$ state, the Zeeman splitting of the $f = 1$, $m_f = \pm 1$ states is given by:

$$E_\pm = E_1 \pm \mu_B B \,, \tag{22.194}$$

and for the coupled $m_f = 0$ states, we must solve the eigenvalue equation:

$$\begin{Vmatrix} E_0 - E & \mu_B B \\ \mu_B B & E_1 - E \end{Vmatrix} = (\, E_0 - E \,)\,(\, E_1 - E \,) - (\mu_B B)^2 = 0 \,, \tag{22.195}$$

which gives two solutions:

$$E'_\pm = (E_0 + E_1)/2 \pm \sqrt{[\,(E_0 - E_1)/2\,]^2 + (\mu_B B)^2} \,. \tag{22.196}$$

We plot the Zeeman hyperfine splitting energies from Eqs. (22.196) and (22.196) in Fig. 22.3. For small values of $\mu_B B$, the hyperfine splitting dominates and the Zeeman splitting is a perturbation, governed by ignoring the off-diagonal terms in Eq. (22.193), whereas for large values of $\mu_B B$, the Zeeman splitting dominates, governed by Eq. (22.192) and the hyperfine splitting is a perturbation.

Exercise 81. Using first order perturbation theory, find the hyperfine splitting when $\mu_B B$ is large, and the states can be described by the vectors: $|\, \ell, m_\ell, s_e, m_{s_e}, s_p, m_{s_p} \,\rangle$. Show that this splitting is just $\Delta/2$, where $\Delta = 1421$ MHz, as indicated in Fig. 22.3.

22.3.10 The Stark effect

The splitting of atomic energy levels as a result of a constant electric field is called the Stark effect. For a constant electric field, the Stark Hamiltonian is given by:

$$H_{\rm S} = +e\,\mathbf{E}\cdot\mathbf{R} \,, \tag{22.197}$$

where e is the magnitude of the charge on the electron, $\mathbf{E}$ the electric field, and $\mathbf{R}$ the center of mass position of the electron. If we put $\mathbf{E} = E_0\,\hat{\mathbf{e}}_z$ in the z-direction, and the electron position in units of the Bohr radius a, the Stark Hamiltonian becomes:

$$H_{\rm S} = e\,a\,E_0\,\bar{R}\,C_{1,0}(\Omega) \,, \tag{22.198}$$

where $\bar{R} = R/a$. Due to the small size of the Bohr radius, it takes an electric field on the order of $10,000$ V/m to produce a splitting energy of 5.29×10^{-5} eV. But this is on the order of the fine structure of hydrogen. So we consider the Stark splitting of the fine structure of hydrogen, where we can describe the states by the coupling: $|\, n, (\ell, s)\, j, m_j \,\rangle$. Matrix elements for the same values of n are given by:

$$\begin{aligned} &\langle\, n, (\ell, s)\, j, m_j \,|\, H_{\rm S} \,|\, n, (\ell', s')\, j', m'_j \,\rangle \\ &\qquad = e\,a\,E_0\,\langle\, n, j \,|\, \bar{R} \,|\, n, j' \,\rangle\,\langle\, n, (\ell, s)\, j, m_j \,|\, C_{1,0}(\Omega) \,|\, n, (\ell', s')\, j', m'_j \,\rangle \,. \end{aligned} \tag{22.199}$$

We worked out this matrix element in Section 21.6.5, and found that for the $n = 1$ level, the splitting vanished. For the $n = 2$ fine structure levels, the only non-zero matrix elements are:

$$\begin{aligned} \langle\, 2s_{1/2}, m \,|\, C_{1,0}(\Omega) \,|\, 2p_{1/2}, m \,\rangle &= -\frac{2}{3}\,m \,, \\ \langle\, 2s_{1/2}, m \,|\, C_{1,0}(\Omega) \,|\, 2p_{3/2}, m \,\rangle &= \frac{1}{3}\,\sqrt{(3/2 - m)(3/2 + m)} = \frac{\sqrt{2}}{3} \,, \end{aligned} \tag{22.200}$$

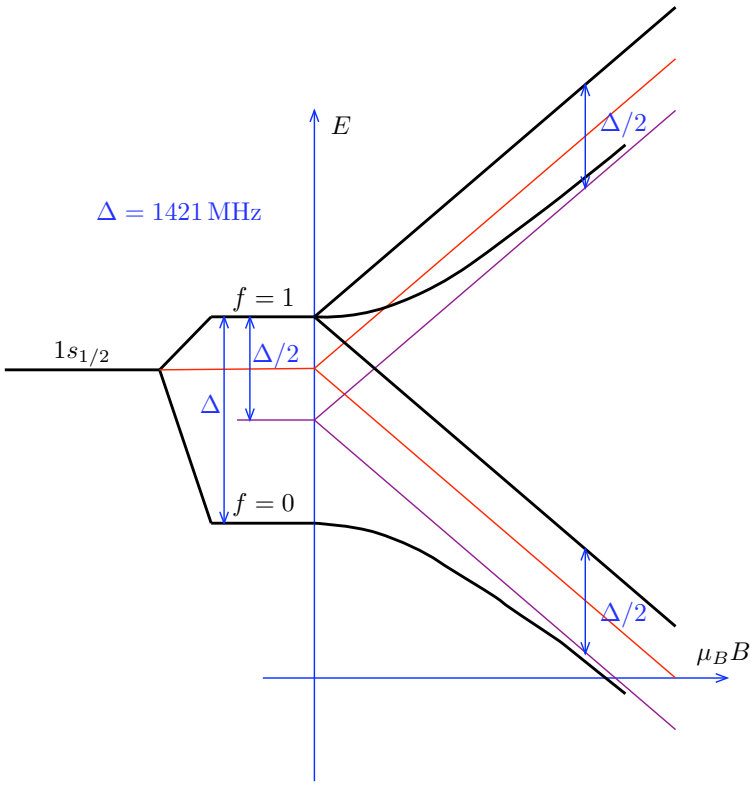

Figure 22.3: Zeeman splitting of the $n = 1$ hyperfine levels of hydrogen as a function of $\mu_B B$ (not to scale).

for $m = \pm 1/2$. Otherwise, the matrix elements vanish. This means that the $2p_{3/2}$ $m = \pm 3/2$ states are not split. Now since the radial integrals are given by [2][p. 239]:

$$\langle\, n, \ell \,|\, \bar{R} \,|\, n, \ell - 1 \,\rangle = -\frac{3}{2}\, n \sqrt{n^2 - \ell^2}\,, \qquad \langle\, 2, 1 \,|\, \bar{R} \,|\, 2, 0 \,\rangle = -3\sqrt{3}\,, \tag{22.201}$$

we find the $m = \pm 1/2$ matrix elements are given by:

$$\begin{aligned} \langle\, 2s_{1/2}, m \,|\, H_{\rm S} \,|\, 2p_{1/2}, m \,\rangle &= 2m\sqrt{3}\,\beta\,, \\ \langle\, 2s_{1/2}, m \,|\, H_{\rm S} \,|\, 2p_{3/2}, m \,\rangle &= -\sqrt{6}\,\beta\,, \end{aligned} \tag{22.202}$$

where $\beta = e\, a\, E_0$. Let us introduce the basis set: $|\,1\,\rangle = |\,2s_{1/2}\,\rangle$, $|\,2\,\rangle = |\,2p_{1/2}\,\rangle$, and $|\,3\,\rangle = |\,2p_{3/2}\,\rangle$. Then the Hamiltonian within the $n = 2$ fine structure levels is given, in matrix form, as:

$$H_m = \begin{pmatrix} -\Delta/3 & 2m\sqrt{3}\,\beta & -\sqrt{6}\,\beta \\ 2m\sqrt{3}\,\beta & -\Delta/3 & 0 \\ -\sqrt{6}\,\beta & 0 & 2\,\Delta/3 \end{pmatrix}\,, \tag{22.203}$$

for $m = \pm 1/2$, and where $\Delta = 4.530 \times 10^{-5}$ eV is the the fine structure splitting between the $2p_{3/2}$ and the $2p_{1/2}$ and $2s_{1/2}$ levels. We have chosen the zero of energy at the "center-of-mass" point so that the sum of

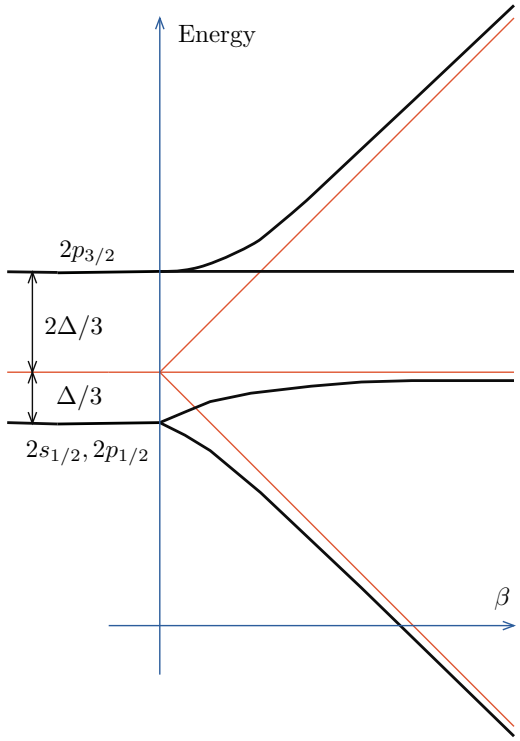

Figure 22.4: Stark splitting of the $n = 2$ fine structure levels of hydrogen as a function of $\beta = e\,a\,E_0$. Δ is the fine structure splitting energy. (not to scale).

the eigenvalues, $\lambda_1 + \lambda_2 + \lambda_3 = 0$, remain zero for all values of β. The energies (λ) and eigenvectors are given by the eigenvalue equation: $H_m \,|\, \lambda \,\rangle = \lambda \,|\, \lambda \,\rangle$, from which we find the cubic equation:

$$\lambda^3 - \left(9\,\beta^2 + \frac{\Delta^2}{3} \right) \lambda - \frac{2}{27}\Delta^3 = 0\,, \tag{22.204}$$

for both values of $m = \pm 1/2$. So the three eigenvalues which are the solutions of Eq. (22.204) are doubly degenerate. The other two energies for the $2p_{3/2}$ state for $m = \pm 3/2$ are not split (in first order perturbation theory), as is the $1s_{1/2}$ ground state. Ignoring the fine structure, the Stark energy splitting is given by: $\Delta E_{m_\ell} = 3\beta\, m_\ell$, where $m_\ell = 0, \pm 1$. Note that for large β, solutions of Eq. (22.204) are given by $\lambda = 0, \pm 3\beta$, in agreement with the result when ignoring the fine structure. The energy levels are sketched in Fig. refem.f:Stark as a function of the electric field strength E_0.

First order perturbation theory gives a vanishing energy shift for the ground state,

$$\Delta E^{(1)}_{1s_{1/2}} = \langle\, 1s_{1/2}, m \,|\, H_{\rm S} \,|\, 1s_{1/2}, m \,\rangle = 0\,. \tag{22.205}$$

So the leading contribution to the ground state energy shift is a second order energy shift, given by:

$$\Delta E^{(2)}_0 = \sum_{\alpha \neq 0} \frac{|\, \langle\, \alpha \,|\, H_{\rm S} \,|\, 0 \,\rangle|^2}{E^{(0)}_0 - E^{(0)}_\alpha}\,. \tag{22.206}$$

Here $|\, 0 \,\rangle$ and $E^{(0)}_0$ are the ground state eigenvector and eigenvalue of the unperturbed hamiltonian H_0. The sum α goes over all excited states of H_0. We can do this sum by a trick invented by Dalgarno and Lewis

[12]. We first note that if we can find an operator $F(\mathbf{R})$ such that:

$$[\,F(\mathbf{R}),H_0\,]\,|\,0\,\rangle = H_S\,|\,0\,\rangle\,, \tag{22.207}$$

then

$$\langle\,\alpha\,|\,[\,F(\mathbf{R}),H_0\,]\,|\,0\,\rangle = \left(\,E_0^{(0)} - E_\alpha^{(0)}\,\right)\langle\,\alpha\,|\,F(\mathbf{R})\,|\,0\,\rangle = \langle\,\alpha\,|\,H_S\,|\,0\,\rangle\,. \tag{22.208}$$

Substituting this into Eq. (22.206), and using completeness of the states of H_0, gives:

$$\begin{aligned}\Delta E_0^{(2)} &= \sum_{\alpha\neq 0}\langle\,0\,|\,H_S\,|\,\alpha\,\rangle\,\langle\,\alpha\,|\,F(\mathbf{R})\,|\,0\,\rangle = \langle\,0\,|\,H_S\,F(\mathbf{R})\,|\,0\,\rangle - \langle\,0\,|\,H_S\,|\,0\,\rangle\,\langle\,0\,|\,F(\mathbf{R})\,|\,0\,\rangle \\ &= \langle\,0\,|\,H_S\,F(\mathbf{R})\,|\,0\,\rangle\,,\end{aligned} \tag{22.209}$$

since $\langle\,0\,|\,H_S\,|\,0\,\rangle = 0$. So the problem reduces to finding a solution for $F(\mathbf{R})$ from Eq. (22.207). It is simplest to do this in coordinate space, where the ground state wave function for H_0 is given by:

$$\psi_0(r) = \frac{1}{\sqrt{\pi a^3}}\,e^{-r/a}\,, \tag{22.210}$$

So Eq. (22.207) becomes:

$$\frac{\hbar^2}{2m}\left[\,\nabla^2\left(F(\mathbf{r})\,e^{-r/a}\right) - F(\mathbf{r})\left(\nabla^2\,e^{-r/a}\right)\right] = eE_0\,r\cos\theta\,e^{-r/a}\,. \tag{22.211}$$

Now we have:

$$\begin{aligned}\nabla^2\left(F(\mathbf{r})\,e^{-r/a}\right) &= \left(\,\nabla^2\left(F(\mathbf{r})\,\right)e^{-r/a} + 2\left(\boldsymbol{\nabla}F(\mathbf{r})\cdot\left(\boldsymbol{\nabla}e^{-r/a}\right) + F(\mathbf{r})\left(\nabla^2\,e^{-r/a}\right)\right.\right. \\ &= \left(\,\nabla^2\left(F(\mathbf{r})\,\right)e^{-r/a} - \frac{2}{a}\,\frac{\partial F(\mathbf{r})}{\partial r}\,e^{-r/a} + F(\mathbf{r})\left(\nabla^2\,e^{-r/a}\right)\right.\,.\end{aligned} \tag{22.212}$$

So (22.211) becomes:

$$\nabla^2\,F(\mathbf{r}) - \frac{2}{a}\,\frac{\partial F(\mathbf{r})}{\partial r} = \frac{2m\,eE_0}{\hbar^2}\,r\cos\theta\,. \tag{22.213}$$

So if we put $F(\mathbf{r}) = f(r)z = f(r)\,r\cos\theta$, then $f(r)$ must satisfy:

$$r\,\frac{\partial^2 f(r)}{\partial r^2} + \left(4 - \frac{2r}{a}\right)\frac{\partial f(r)}{\partial r} - \frac{2}{a}\,\frac{\partial f(r)}{\partial r} = 2\kappa\,r\,, \qquad \kappa = \frac{m\,eE_0}{\hbar^2}\,. \tag{22.214}$$

A particular solution of this differential equation is:

$$f(r) = -\kappa a\left(\frac{r}{2} + a\right). \tag{22.215}$$

Putting this result into Eq. (22.209), and taking expectation values in the ground state gives:

$$\begin{aligned}\Delta E_0^{(2)} &= -\kappa a\,eE_0\,\langle\,z^2\,(r/2+a)\,\rangle = -\frac{\kappa a eE_0}{3}\left\{\,\langle\,r^3\,\rangle/2 + a\,\langle\,r^2\,\rangle\,\right\} \\ &= -\frac{a^3\,E_0^2}{3}\left\{\,\langle\,\bar{r}^3\,\rangle/2 + \langle\,\bar{r}^2\,\rangle\,\right\} = -\frac{9}{4}\,a^3\,E_0^2\,.\end{aligned} \tag{22.216}$$

The fact that the second order perturbation theory for the Stark effect can be summed for hydrogen came as a surprise in the 1955. We should, perhaps, consider this a result of the $SO(4,2)$ symmetry of the unperturbed Hamiltonian for hydrogen.

22.4 Atomic radiation

In this section, we discuss the interaction of time dependent electromagnetic radiation with an electron in an atom. This is called a *semi-classical* approximation since the motion of the electron is treated quantum mechanically but the electromagnetic radiation field is treated classically. As long as we can ignore individual photon effects of the electromagnetic radiation field, this is a good approximation.

There are two phenomena we wish to discuss in a general way: (1) the absorption of radiation by an electron in an atom, and (2) the production of radiation by an electron undergoing oscillations between energy levels in an atom.

22.4.1 Atomic transitions

In this section, we discuss transitions of electrons between energy levels in atoms caused by electromagnetic radiation.

22.4.2 The photoelectric effect

The photoelectric effect is the ejection of electrons by atoms caused by electromagnetic radiation.

22.4.3 Resonance fluorescence

Resonance fluorescence is often described as absorption and re-radiation of electromagnetic

22.5 Magnetic flux quantization and the Aharonov-Bohm effect

One of the striking properties of superconducting rings is that magnetic flux can get trapped in the ring in quantized amounts. Experiments with superconductors were first done by Dever and Fairbanks in the 1960's and the theory worked out by Lee and Byers. Similar experiments with free electrons passing on each side of magnetic needles go under the name of the Aharonov-Bohm effect. In both of these experiments, there is an effect of a magnetic field on particles in regions where there *is* no magnetic field present; however the region of space is not simply connected but contains "holes." Both of these surprising results happen because of quantum mechanics, there is no classical analogue. We first discuss the quantized flux experiments.

22.5.1 Quantized flux

In the superconductor quantized flux experiments, a ring is placed in a magnetic field, and the temperature brought down below the superconducting temperature. The magnetic field is then turned off thus trapping flux through the hole in the the superconducting ring. The current flowing in the ring maintains the trapped flux indefinitely. It is this trapped flux that is observed to be in quantized amounts. Let us analyze this experiment.

The first thing we have to understand is that the magnetic field $\mathbf{B}$ is completely excluded from the superconductor so that there is no Lorentz force acting on the electrons. However, there is a vector potential $\mathbf{A}$ inside the superconductor, which, in the Schrödinger picture, will have an effect on the wave function. For $r \geq a$, where where a is the inside radius of the superconducting ring, the magnetic flux is given by:

$$\Phi_B = \int \mathbf{B} \cdot \mathrm{d}\mathbf{S} = \oint \mathbf{A} \cdot \mathrm{dl}\,,$$

So for a uniform field $\mathbf{B}$ in the z-direction, the magnetic flus is given by:

$$\Phi_B = \pi a^2\, B = 2\pi r\, A\,, \qquad \text{for } r > a,$$

In cylindrical coordinates, the vector potential is given by:

$$\mathbf{A}(r) = \frac{\Phi_B}{2\pi r}\,\hat{\mathbf{e}}_\phi\,, \qquad \text{for } r \geq a.$$

Now we can, by a gauge transformation, remove this vector potential completely from the problem. That is, if we choose

$$\Lambda(\mathbf{r}) = \frac{\Phi_B}{2\pi}\,\phi\,,$$

then the scalar potential remains the same, but the new vector potential vanishes. In cylindrical coordinates,

$$\boldsymbol{\nabla} = \hat{\mathbf{e}}_r\,\frac{\partial}{\partial r} + \hat{\mathbf{e}}_\phi\,\frac{1}{r}\,\frac{\partial}{\partial\phi} + \hat{\mathbf{e}}_z\,\frac{\partial}{\partial z}\,,$$

so:

$$\mathbf{A}'(\mathbf{r}) = \mathbf{A}(\mathbf{r}) - \boldsymbol{\nabla}\Lambda(\mathbf{r}) = 0\,.$$

Thus the wave function is given by:

$$\psi(\mathbf{r},t) = \exp\Big\{-\frac{iq}{\hbar c}\,\frac{\Phi_B}{2\pi}\,\phi\Big\}\,\psi'(\mathbf{r},t)\,.$$

where $\psi'(\mathbf{r},t)$ is the solution of Schrödinger's equation with no vector potential, and is single-valued about the hole in the superconductor. But $\psi'(\mathbf{r},t)$ must *also* be single-valued about the hole, so we must require:

$$\frac{q}{\hbar c}\,\Phi_B = 2\pi\,n\,,$$

and we find that the flux Φ_B is restricted to the quantized values:

$$\Phi_B = \frac{2\pi\hbar c}{q}\,n = \frac{\pi\hbar c}{e}\,n\,. \tag{22.217}$$

Here, we have set $q = 2e$ since the charge carriers in superconductors are electrons and holes pairs. The essential feature of this analysis is that even though the vector potential for this problem can be removed by a gauge transformation, the gauge potential depends on position and for problems where there are "holes" in the allowed region for the electron, this can lead to physical consequences.

22.5.2 The Aharonov-Bohm effect

When an electron beam is required to pass on both sides of a region of magnetic flux contained in a tube of radius a, the diffraction pattern produced also exhibites a phase shift due to the magnetic flux in the hole, even though there is no magnetic field in the region where the particle is. This is called the Aharonov-Bohm effect. Let us analyze this experiment. The wave function for an electron which can take the two paths shown in the figure is the sum of the solutions of Schrödinger's equation for the two paths, which, for simplicity, we take to be plane waves given by:

$$\psi_k(\mathbf{r},t) = C_1\,e^{i(kx_1-\omega_k\,t)} + C_2\,e^{i(kx_2-\omega_k\,t)}\,, \tag{22.218}$$

with $E = \hbar^2k^2/(2m) = \hbar\,\omega_k$, and where x_1 and x_2 are the two paths shown in the Figure. The wave function, however, must be single-valued around a source of flux.

Finish this derivation!

22.6 Magnetic monopoles

Maxwell's equations does not preclude the possibility of the existance of magnetic monopoles. If magnetic charge and current exists, Maxwell's equations reads:

$$\boldsymbol{\nabla}\cdot\mathbf{E} = 4\pi\rho_e\,, \qquad \boldsymbol{\nabla}\times\mathbf{B} = \frac{1}{c}\frac{\partial\mathbf{E}}{\partial t} + \frac{4\pi}{c}\mathbf{J}_e$$
$$\boldsymbol{\nabla}\cdot\mathbf{B} = 4\pi\rho_m\,, \qquad -\boldsymbol{\nabla}\times\mathbf{E} = \frac{1}{c}\frac{\partial\mathbf{B}}{\partial t} + \frac{4\pi}{c}\mathbf{J}_m$$

These equations are invariant under a "duality" transformation in which electric and magnetic fields, charges, and currents are rotated in an abstract vector space. Consider the duality transformation, given by:

$$\begin{pmatrix}\mathbf{E}'\\ \mathbf{B}'\end{pmatrix} = \begin{pmatrix}\cos\chi & \sin\chi\\ -\sin\chi & \cos\chi\end{pmatrix}\begin{pmatrix}\mathbf{E}\\ \mathbf{B}\end{pmatrix}\,,$$

with similar expressions for rotation of the coordinates. Here χ represents the "rotation" of $\mathbf{E}$ into $\mathbf{B}$. Both the electric and magnetic sources here obey conservation equations.

Now according to these equations, a single static magnetic "charge" q_m, can generate a coulomb-like magnetic field. Thus we find:

$$\mathbf{B} = \frac{q_m}{r^2}\,\hat{\mathbf{e}}_r\,.$$

We can still define the vector potentials by

$$\mathbf{E}(\mathbf{r},t) = -\boldsymbol{\nabla}\phi(\mathbf{r},t) - \frac{1}{c}\frac{\partial\mathbf{A}(\mathbf{r},t)}{\partial t}\,,$$
$$\mathbf{B}(\mathbf{r},t) = \boldsymbol{\nabla}\times\mathbf{A}(\mathbf{r},t)\,.$$

However for monopoles, the second of these contradicts Maxwell's equations if $\mathbf{A}(\mathbf{r})$ is a single valued vector function of $\mathbf{r}$. We can get around this problem by constructing a double valued vector potential, valid in two regions of space. By examining the form of the curl in spherical coordinates, we see that we can obtain the correct $\mathbf{B}$ from the two expressions,

$$\mathbf{A}^I(\theta) = +\left\{\frac{q_m(1-\cos\theta)}{r\sin\theta}\right\}\hat{\mathbf{e}}_\phi\,, \qquad \text{for } 0\le\theta\le\pi-\epsilon,$$
$$\mathbf{A}^{II}(\theta) = -\left\{\frac{q_m(1+\cos\theta)}{r\sin\theta}\right\}\hat{\mathbf{e}}_\phi\,, \qquad \text{for } \epsilon\le\theta\le\pi.$$

However in the overlap region, $\epsilon\le\theta\le\pi-\epsilon$, the two expressions for $\mathbf{A}$ give the same $\mathbf{B}$ field, and therefore must be related by a gauge transformation,

$$\mathbf{A}^{II}(\mathbf{r}) = \mathbf{A}^I(\mathbf{r}) - \boldsymbol{\nabla}\Lambda(\mathbf{r})\,.$$

Using the expression for the gradient in spherical coordinates, we find that the gauge field $\Lambda(\phi)$ is given by:

$$\Lambda(\phi) = 2q_m\,\phi\,. \tag{22.219}$$

Now suppose a second particle has an electric charge q_e. The Lagrangian for the interaction of this electrically charged particle with the vector potential associated with the magnetic monopole is given by:

$$L = q_e\,\mathbf{v}\cdot\mathbf{A}(\mathbf{r},t)/c$$

At first, it might be surprising that there is an interaction between a magnetic monopole charge and an electric monopole charge. This is because in the Schrödinger representation, the electric charge interects by means of the vector potential, and the vector potential comes from the magnetic monopole. However in the

overlap region where the two vector potentials are related by a gauge transformation, the wave functions in the Schrödinger picture in the two gauges are given by:

$$\psi^{II}(\mathbf{r},t) = \exp\Big\{ -i\frac{2q_eq_m}{\hbar c}\,\phi \Big\}\,\psi^{I}(\mathbf{r},t)\,.$$

Therefore, since both ψ^I and ψ^{II} must be single valued, we must have

$$\frac{2q_eq_m}{\hbar c} = n\,, \qquad \text{for } n = \pm1, \pm2, \ldots,. \tag{22.220}$$

This means that if a monopole exists, the electric charge must be quantized:

$$\frac{q_e}{e} = \frac{\hbar c}{2e^2}\frac{e}{q_m}\,n = \frac{1}{2\alpha}\frac{e}{q_m}\,n\,,$$

where $\alpha = e^2/\hbar c = 1/137$ is the fine structure constant. As a result of this analysis, we are led to say that if electric charge is quantized in units of e, and if this is due to the existence of magnetic monopoles, then the magnetic charge of the monopole must be such that $2\alpha\, q_m = e$. (Quarks, we think, have fractional charges!)

In our arguement, we only had to suppose that somewhere in the universe there existed a magnetic monopole that can interact with any charged particle. Then gauge consistency required that the electric charge is quantized. This remarkable results was first given by Dirac in 1931. Generally duality theories in particle physics predict the existance of magnetic monopoles; since they haven't been observed, either such theories are incorrect or the magnetic monopoles have been banished to the edges of the universe.

[This section needs further work as it is hard to follow. I am not sure it is worth it to include since this topic doesn't seem to be very hot now-a-days, and no magnetic monopoles have ever been found.]

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).

[2] H. A. Bethe and E. E. Salpeter, *Quantum mechanics of one- and two-electron atoms* (Springer-Verlag (Berlin), Academic Press, New York. NY, 1957).

ANNOTATION: This book is Volume XXXV of the Springer-Verlag Encyclopedia of Physics, and is based on the earlier 1932 Geiger-Scheel Handbuch der Physik article by H. A. Bethe. Bethe refers to this as a "low-brow" book, devoted to practical calculations. It contains many useful results in the atomic physics of hydrogen and helium.

[3] E. Schrödinger, "The relationship between the quantum mechanics of Heisenberg, Born and Jordan, and that of Schrödinger," Ann. Phys. **79**, 361 (1926).

[4] W. Pauli, "Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik," Zeit. für Physik **36**, 336 (1926).

[5] C. Runge, *Vektoranalysis*, volume 1 (S. Hirzel, Leipzig, 1919). P. 70.

[6] W. Lenz, "Über den Bewegungsverlauf und die Quantenzustände der gestörten Keplerbewegung," Zeit. für Physik **24**, 197 (1924).

[7] V. Barger and M. Olsson, *Classical mechanics: a modern perspective* (McGraw-Hill, New York, 1995), second edition.

[8] E. Schrödinger, "A method of determining quantum mechanical eigenvalues and eigenfunctions," Proc. Roy. Irish Acad. A **46**, 9 (1940).

[9] L. Infeld and T. E. Hull, "The factorization method," Rev. Mod. Phys. **23**, 21 (1951).

[10] K. T. Hecht, *Quantum Mechanics*, Graduate texts in contemporary physics (Springer, New York, 2000).

[11] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, NY, 1964).

[12] A. Dalgarno and J. T. Lewis, "The exact calculation of long-range forces between atoms by perturbation theory," Proc. Roy. Soc. London **233**, 70 (1955).

Chapter 23

Scattering theory

The scattering of particles has been one of the important tools to study the fundamental forces in nature. Thus it is important to understand how these experiments are carried out and what information can be obtained from them. Of course, scattering takes place naturally in physical processes, and so we need to know what happens during these events. In this chapter, we illustrate scattering by detailed analysis of several examples.

23.1 Propagator theory

23.1.1 Free particle Green function in one dimension

In example 22 on page 40 and in remark 11, we found the Dirac bracket $\langle\, q,t \,|\, q't' \,\rangle$ in the Heisenberg representation for a free particle in the coordinate representation:

$$\langle\, q,t \,|\, q't' \,\rangle = \sqrt{\frac{m}{2\pi i\hbar\,(t-t')}}\,\exp\left\{\frac{i}{\hbar}\,\frac{m}{2}\,\frac{(q-q')^2}{(t-t')}\right\}\,. \tag{23.1}$$

We can use this bracket to find the free particle wave function at time t, given the free particle wave function at time t'. From the properties of the Dirac brackets, we have:

$$\begin{aligned}\psi(q,t) = \langle\, q,t \,|\, \psi \,\rangle &= \int_{-\infty}^{+\infty} \mathrm{d}q'\,\langle\, q,t \,|\, q',t' \,\rangle\,\langle\, q',t' \,|\, \psi \,\rangle \\ &= \sqrt{\frac{m}{2\pi i\hbar\,t}}\,\int_{-\infty}^{+\infty}\mathrm{d}q'\,\exp\left\{\frac{i}{\hbar}\,\frac{m}{2}\,\frac{(q-q')^2}{(t-t')}\right\}\,\psi(q',t')\,.\end{aligned} \tag{23.2}$$

Note that there is an integration only over q', not t'. So the bracket $\langle\, q,t \,|\, q't' \,\rangle$ is a Green function for a free particle. Let us define a Green function $G(q,t;q't')$ by the equation:

$$\left\{\,i\hbar\,\frac{\partial}{\partial t} - \frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial q^2}\,\right\}\,G(q,t;q',t') = \frac{\hbar}{i}\,\delta(q-q')\,\delta(t-t')\,. \tag{23.3}$$

The solution of this equation can be found by a double fourier transform. We let:

$$G(q,t;q',t') = \int_{-\infty}^{+\infty}\frac{\mathrm{d}p}{2\pi\hbar}\int_{-\infty}^{+\infty}\frac{\mathrm{d}E}{2\pi\hbar}\,\tilde{G}(p,E)\,e^{i[\,p\,(q-q')-E\,(t-t')]/\hbar}\,. \tag{23.4}$$

Then since

$$\int_{-\infty}^{+\infty}\frac{\mathrm{d}p}{2\pi\hbar}\int_{-\infty}^{+\infty}\frac{\mathrm{d}E}{2\pi\hbar}\,e^{i[\,p\,(q-q')-E\,(t-t')]/\hbar} = \delta(q-q')\,\delta(t-t')\,, \tag{23.5}$$

we find that $\tilde{G}(p,E)$ is given by the solution of:

$$\left\{ E - \frac{p^2}{2m} \right\} \tilde{G}(p,E) = \frac{\hbar}{i} \,. \tag{23.6}$$

There is no solution to this equation when $E = p^2/(2m)$. However we can find solutions by first allowing E to be a complex variable. Then we introduce a small imaginary part $\pm i\epsilon$ to E, and take the limit $\epsilon \to 0$ after the integration. This gives two solutions, called the retarded $(+)$ and advanced $(-)$ solutions:

$$\tilde{G}^{(\pm)}(p,E) = \frac{\hbar}{i} \frac{1}{E - p^2/(2m) \pm i\epsilon} \,. \tag{23.7}$$

For the retarded Green function $\tilde{G}^{(+)}(q,t;q',t')$, we find:

$$\begin{aligned} G^{(+)}(q,t;q',t') &= \frac{\hbar}{i} \int_{-\infty}^{+\infty} \frac{\mathrm{d}p}{2\pi\hbar} \int_{-\infty}^{+\infty} \frac{\mathrm{d}E}{2\pi\hbar} \frac{e^{i[\,p\,(q-q') - E\,(t-t')]/\hbar}}{E - p^2/(2m) + i\epsilon} \\ &= \Theta(t-t') \int_{-\infty}^{+\infty} \frac{\mathrm{d}p}{2\pi\hbar} \exp\left\{ \frac{i}{\hbar} \left[p\,(q-q') - \frac{p^2}{2m}\,(t-t') \right] \right\} \\ &= \Theta(t-t')\, \langle\, q,t \,|\, q',t' \,\rangle \,, \end{aligned} \tag{23.8}$$

where $\langle\, q,t \,|\, q',t' \,\rangle$ is given by Eq. (23.1). For the advanced Green function $G^{(-)}(q,t;q',t')$, we find:

$$\begin{aligned} G^{(-)}(q,t;q',t') &= \frac{\hbar}{i} \int_{-\infty}^{+\infty} \frac{\mathrm{d}p}{2\pi\hbar} \int_{-\infty}^{+\infty} \frac{\mathrm{d}E}{2\pi\hbar} \frac{e^{i[\,p\,(q-q') - E\,(t-t')]/\hbar}}{E - p^2/(2m) - i\epsilon} \\ &= -\Theta(t'-t) \int_{-\infty}^{+\infty} \frac{\mathrm{d}p}{2\pi\hbar} \exp\left\{ \frac{i}{\hbar} \left[p\,(q-q') - \frac{p^2}{2m}\,(t-t') \right] \right\} \\ &= -\Theta(t'-t)\, \langle\, q,t \,|\, q',t' \,\rangle \,. \end{aligned} \tag{23.9}$$

So we find that:

$$\langle\, q,t \,|\, q',t' \,\rangle = \theta(t-t')\, G^{(+)}(q,t;q',t') - \theta(t'-t)\, G^{(-)}(q,t;q',t') = \begin{cases} +G^{(+)}(q,t;q',t') & \text{for } t > t', \\ -G^{(-)}(q,t;q',t') & \text{for } t < t'. \end{cases} \tag{23.10}$$

23.2 S-matrix theory

23.3 Scattering from a fixed potential

23.4 Two particle scattering

We begin by studying the scattering between two distinguishable, spinless particles of mass m_1 and m_2. We suppose, for example, that particle 1 is incident along the z direction in the laboratory with kinetic energy E_1 to a target particle 2 at rest at the origin, as illustrated in the figure. We assume that the interaction between the two particles can be represented by a potential that depends only on the magnitude of the distance between them. For this reason, the problem is best solved in the center of mass coordinate system. The relation between these systems is given by:

$$\begin{aligned} \mathbf{R} &= \frac{m_1}{M}\mathbf{r}_1 + \frac{m_2}{M}\mathbf{r}_2 \,, & \mathbf{r}_1 &= \mathbf{R} + \frac{\mu}{m_1}\mathbf{r} \,, \\ \mathbf{r} &= \mathbf{r}_1 - \mathbf{r}_2 \,, & \mathbf{r}_2 &= \mathbf{R} - \frac{\mu}{m_2}\mathbf{r} \,, \end{aligned}$$

or where the total mass $M = m_1 + m_2$ and the reduced mass $\mu = m_1 m_2/M$. We have also define:

$$\boldsymbol{\nabla}_R = \boldsymbol{\nabla}_1 + \boldsymbol{\nabla}_2 \,, \qquad \boldsymbol{\nabla}_1 = \frac{m_1}{M}\boldsymbol{\nabla}_R + \boldsymbol{\nabla}_r \,,$$
$$\boldsymbol{\nabla}_r = \frac{\mu}{m_1}\boldsymbol{\nabla}_1 - \frac{\mu}{m_2}\boldsymbol{\nabla}_2 \,, \qquad \boldsymbol{\nabla}_2 = \frac{m_2}{M}\boldsymbol{\nabla}_R - \boldsymbol{\nabla}_r \,.$$

With these definitions, we can show that

$$\mathbf{r}_1 \cdot \boldsymbol{\nabla}_1 + \mathbf{r}_2 \cdot \boldsymbol{\nabla}_2 = \mathbf{R} \cdot \boldsymbol{\nabla}_R + \mathbf{r} \cdot \boldsymbol{\nabla}_r$$
$$\frac{\hbar^2}{2m_1}\boldsymbol{\nabla}_1^2 + \frac{\hbar^2}{2m_2}\boldsymbol{\nabla}_2^2 = \frac{\hbar^2}{2M}\boldsymbol{\nabla}_R^2 + \frac{\hbar^2}{2\mu}\boldsymbol{\nabla}_r^2 \,.$$

So it is useful the define total and relative wave numbers in the same way as the nabla operators:

$$\mathbf{K} = \mathbf{k}_1 + \mathbf{r}_2 \,, \qquad \mathbf{k}_1 = \frac{m_1}{M}\mathbf{K} + \mathbf{k} \,,$$
$$\mathbf{k} = \frac{\mu}{m_1}\mathbf{k}_1 - \frac{\mu}{m_2}\mathbf{k}_2 \,, \qquad \mathbf{k}_2 = \frac{m_2}{M}\mathbf{K} - \mathbf{k} \,.$$

Then we can prove that

$$\mathbf{k}_1 \cdot \mathbf{r}_1 + \mathbf{k}_2 \cdot \mathbf{r}_2 = \mathbf{K} \cdot \mathbf{R} + \mathbf{k} \cdot \mathbf{r} \,,$$
$$E = \frac{\hbar^2 k_1^2}{2m_1} + \frac{\hbar^2 k_1^2}{2m_1} = \frac{\hbar^2 K^2}{2M} + \frac{\hbar^2 k^2}{2\mu} \,. \tag{23.11}$$

The Hamiltonian for this problem in the two systems is given by:

$$H = -\frac{\hbar^2}{2m_1}\boldsymbol{\nabla}_1^2 - \frac{\hbar^2}{2m_2}\boldsymbol{\nabla}_2^2 + V(|\mathbf{r}_1 - \mathbf{r}_2|) \,, \tag{23.12}$$
$$= -\frac{\hbar^2}{2M}\boldsymbol{\nabla}_R^2 - \frac{\hbar^2}{2\mu}\boldsymbol{\nabla}_r^2 + V(r) \,. \tag{23.13}$$

So we can separate variables for the wave function for Schrödinger's time independent equation in the relative and center of mass system by the ansaz:

$$\psi_{\mathbf{K},\mathbf{k}}(\mathbf{R},\mathbf{r}) = e^{i\mathbf{K}\cdot\mathbf{R}}\,\psi_{\mathbf{k}}(\mathbf{r}) \,, \tag{23.14}$$

where $\psi_{\mathbf{k}}(\mathbf{r})$ satisfies:

$$\left\{-\frac{\hbar^2}{2\mu}\boldsymbol{\nabla}_r^2 + V(r)\right\}\psi_{\mathbf{k}}(\mathbf{r}) = \frac{\hbar^2 k^2}{2\mu}\,\psi_{\mathbf{k}}(\mathbf{r}) \,. \tag{23.15}$$

Now as $r \to \infty$, we assume that $V(r) \to 0$ sufficiently rapidly, so that solutions of (23.15) become solutions for a free particle. We want to require this to be the sum of an incident wave and a scattered wave. Thus we require solutions of (23.15) to have the asymptotic form,

$$\psi_{\mathbf{k}}(r,\theta,\phi) \sim e^{i\mathbf{k}\cdot\mathbf{r}} + f_k(\theta)\frac{e^{ikr}}{r} \,. \tag{23.16}$$

Here $\cos\theta = \hat{\mathbf{r}} \cdot \hat{\mathbf{k}}$, the angle between the incident particle in the relative system (this is the z-axis in our coordinate system) and the scattering direction $\hat{\mathbf{r}}$. The total energy is given by (23.11), which, for a given $\mathbf{k}_1$ and $\mathbf{k}_2$ define both $\mathbf{K}$ and $\mathbf{k}$. Requirement (23.16) gives for the incident and scattering flux:

$$\mathbf{j}_{\text{inc}} = \left(\frac{\hbar\mathbf{k}}{m}\right) \,,$$
$$\mathbf{j}_{\text{scat}} \sim \left(\frac{\hbar k}{m}\right)\frac{|f_k(\theta)|^2}{r^2}\hat{\mathbf{r}} \,,$$

and thus the differential scattering cross section is given by:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{4\pi r^2\, j_{\mathrm{scat}}}{j_{\mathrm{inc}}} = |f_k(\theta)|^2 \,.$$

So the physics is contained in the scattering amplitude $f_k(\theta)$, and we turn now to its calculation in terms of the potential function $V(r)$ in Schrödinger's equation (23.15). We will show that it depends only on k and $\cos\theta$.

Since the potential in (23.15) depends only on r, we can separate variables, and write

$$\psi_{\mathbf{k}}(r,\theta,\phi) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} R_{k\ell}(r)\, Y_{\ell m}(\theta,\phi) \,, \tag{23.17}$$

where $Y_{\ell m}(\hat{\mathbf{r}})$ are the spherical harmonics and $R_{k\ell}(r)$ satisfies the radial equation,

$$\left\{-\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\ell(\ell+1)}{r^2} + w(r)\right\} R_{k\ell}(r) = k^2\, R_{k\ell}(r) \,, \tag{23.18}$$

where $w(r)$ is defined by:

$$V(r) = \frac{\hbar^2}{2\mu}\, w(r) \,. \tag{23.19}$$

The solutions $R_{k\ell}(r)$ of (23.18) are independent of m.

Example 41. For a square well potential, given by:

$$V(r) = \begin{cases} -V_0 = -\dfrac{\hbar^2}{2\mu}\, w_0 \,, & \text{for } r \le a, \\ 0 \,, & \text{for } r > a, \end{cases} \tag{23.20}$$

solutions of (23.18), which are regular at the origin, are given by:

$$R_{k\ell}(r) = \begin{cases} D_\ell(k)\, j_\ell(\kappa r) \,, & \text{for } r \le a, \\ A_\ell(k)\, j_\ell(kr) + B_\ell(k)\, n_\ell(kr) \,, & \text{for } r > a, \end{cases} \tag{23.21}$$

where $\kappa = \sqrt{k^2 + w_0}$, and j_ℓ and n_ℓ are the regular and irregular (real) spherical Bessel functions. The coefficients $A_\ell(k)$, $B_\ell(k)$, and $D_\ell(k)$ have to be picked so that the solution $R_{k\ell}(r)$ is continuous and has continuous derivatives at $r = a$. This gives the requirements,

$$D_\ell(k)\, j_\ell(\kappa a) = A_\ell(k)\, j_\ell(ka) + B_\ell(k)\, n_\ell(ka) \,, \tag{23.22}$$

$$\kappa\, D_\ell(k)\, j'_\ell(\kappa a) = k\, A_\ell(k)\, j'_\ell(ka) + k\, B_\ell(k)\, n'_\ell(ka) \,. \tag{23.23}$$

So if we put

$$\begin{aligned} A_\ell(k) &= C_\ell(k)\, \cos\delta_\ell(k) \,, \\ B_\ell(k) &= -C_\ell(k)\, \sin\delta_\ell(k) \,, \end{aligned} \tag{23.24}$$

which defines the phase shifts $\delta_\ell(k)$, then solutions of (23.22) and (23.23) are given by:

$$\tan\delta_\ell(k) = \frac{\kappa j'_\ell(\kappa a) j_\ell(ka) - k j_\ell(\kappa a) j'_\ell(ka)}{\kappa j'_\ell(\kappa a) n_\ell(ka) - k j_\ell(\kappa a) n'_\ell(ka)} \,, \tag{23.25}$$

$$D_\ell(k) = C_\ell(k)\, k\, [\, j'_\ell(ka) n_\ell(ka) - j'_\ell(ka) n_\ell(ka)\,] \,/\, N_\ell(k) \,, \tag{23.26}$$

$$N_\ell(k) = \sqrt{X_\ell^2(k) + Y_\ell^2(k)} \,, \tag{23.27}$$

$$X_\ell(k) = \kappa j'_\ell(\kappa a) j_\ell(ka) - k j_\ell(\kappa a) j'_\ell(ka) \,, \tag{23.28}$$

$$Y_\ell(k) = \kappa j'_\ell(\kappa a) n_\ell(ka) - k j_\ell(\kappa a) n'_\ell(ka) \,, \tag{23.29}$$

which gives $D_\ell(k)$ in terms of $C_\ell(k)$, which will be fixed by the asymptotic conditions below.

As long as $V(r) \to 0$ as $r \to \infty$ sufficiently rapidly, the radial solution is given by a linear combination of the free particle solutions. Thus for any potential that satisfies this criterion, as $r \to \infty$, $R_{k\ell}(r)$ has the asymptotic form,

$$\begin{aligned} R_{k\ell}(r) &\sim A_\ell(k)\, j_\ell(kr) + B_\ell(k)\, n_\ell(kr)\,, \\ &= C_\ell(k)\, \{\cos\delta_\ell(k)\, j_\ell(kr) - \sin\delta_\ell(k)\, n_\ell(kr)\}\,, \\ &\sim C_\ell(k)\, \{\cos\delta_\ell(k)\, \sin(kr-\ell\pi/2) + \sin\delta_\ell(k)\, \cos(kr-\ell\pi/2)\}/kr\,, \\ &= C_\ell(k)\, \{\sin(kr-\ell\pi/2+\delta_\ell(k))\}/kr\,, \end{aligned}$$

where we have used the asymptotic forms,

$$j_\ell(kr) \sim +\sin(kr-\ell\pi/2)/(kr)\,, \quad \text{as } r\to\infty. \tag{23.30}$$

$$n_\ell(kr) \sim -\cos(kr-\ell\pi/2)/(kr)\,, \quad \text{as } r\to\infty. \tag{23.31}$$

Therefore as $r \to \infty$, the solution to Schrödinger's equation is given by:

$$\begin{aligned} \psi_k(r,\theta,\phi) &\sim \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} C_\ell(k)\, \frac{\sin(kr-\ell\pi/2+\delta_\ell(k))}{kr}\, Y_{\ell m}(\theta,\phi) \\ &= \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} C_\ell(k)\, \left\{\frac{e^{i[kr-\ell\pi/2+\delta_\ell(k)]} - e^{-i[kr-\ell\pi/2+\delta_\ell(k)]}}{2ikr}\right\} Y_{\ell m}(\theta,\phi)\,, \end{aligned} \tag{23.32}$$

which contains both ingoing and outgoing waves. On the other hand, from the asymptotic form of the wave function, we find:

$$\begin{aligned} \psi_k(r,\theta) &\sim e^{ikz} + f_k(\theta)\frac{e^{ikr}}{r} \\ &= 4\pi\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} Y^*_{\ell m}(\hat{\mathbf{k}})\, i^\ell\, j_\ell(kr)\, Y_{\ell m}(\theta,\phi) + f_k(\theta)\frac{e^{ikr}}{r} \\ &\sim 4\pi\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} Y^*_{\ell m}(\hat{\mathbf{k}})\, i^\ell \left\{\frac{\sin(kr-\ell\pi/2)}{kr}\right\} Y_{\ell m}(\theta,\phi) + f_k(\theta)\frac{e^{ikr}}{r} \\ &= \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} Y^*_{\ell m}(\hat{\mathbf{k}})\, i^\ell \left\{\frac{e^{i[kr-\ell\pi/2]} - e^{-i[kr-\ell\pi/2]}}{2ikr}\right\} Y_{\ell m}(\theta,\phi) \\ &\quad + f_k(\theta)\frac{e^{ikr}}{r} \end{aligned} \tag{23.33}$$

Here, we have used the relation

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{\ell=0}^{\infty}(2\ell+1)\, i^\ell j_\ell(kr) P_\ell(\hat{\mathbf{k}}\cdot\hat{\mathbf{r}})\,, \tag{23.34}$$

$$= 4\pi\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} i^\ell j_\ell(kr) Y^*_{\ell m}(\hat{\mathbf{k}}) Y_{\ell m}(\hat{\mathbf{r}})\,. \tag{23.35}$$

Comparing coefficients of the ingoing wave (e^{-ikr}) in equations (23.32) and(23.33), we find:

$$C_\ell(\mathbf{k}) = 4\pi Y^*_{\ell m}(\hat{\mathbf{k}})\, i^\ell\, e^{i\delta_\ell(k)}\,. \tag{23.36}$$

This gives, for the coefficients of the outgoing wave (e^{ikr}),

$$f_k(\theta) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{e^{2i\delta_\ell(k)} - 1}{2ik} Y^*_{\ell m}(\hat{\mathbf{k}}) Y_{\ell m}(\hat{\mathbf{r}}) \,,$$
$$= \frac{1}{k} \sum_{\ell=0}^{\infty} (2\ell + 1)\, e^{i\delta_\ell(k)} \sin\delta_\ell(k)\, P_\ell(\cos\theta) \,. \tag{23.37}$$

So the scattering amplitude depends only on $\cos\theta$, as expected. Eq. (23.37) is exact. We only need the phase shifts $\delta_\ell(k)$, which must be found by solving the radial Schrödinger equation for a given potential.

Note the different forms for the partial scattering amplitude:

$$f_\ell(k) = \frac{e^{i\delta_\ell(k)} \sin\delta_\ell(k)}{k} = \frac{1}{k(\cot\delta_\ell(k) - i)} \,, \tag{23.38}$$

and that $|2ikf_\ell(k) + 1| = 1$ (It's always on the unit circle!). At points where $\delta_\ell(k)$ is a multiple of $\pi/2$, the scattering amplitude is maximal.

Example 42 (The optical theorem)**.** Show that:

$$\sigma = \int |f_k(\theta)|^2 \, \mathrm{d}\Omega = \frac{4\pi}{k^2} \sum_{\ell=0}^{\infty} (2\ell + 1) \sin^2\delta_\ell(k) = \frac{4\pi}{k} \mathrm{Im}\{f_k(0)\} \,. \tag{23.39}$$

The optical theorem states that the total probability for scattering of a particle is equal to the loss of probability from the incident beam. (Show this also.)

From classical arguements, we can understand how many phase shifts must be found for a given incident wave number k. In classical scattering, the angular momentum is related to the impact parameter by $\ell = kb < ka$, where a is the range of the potential (the radius, for the case of a square well). So if $ka \sim 10$, about 10 phase shifts must be calculated to obtain an accurate result for the scattering amplitude.

Example 43. Show that for a square well, the phase shifts for $\ell \gg ka$ can be neglected.

Example 44. For low incident energy, $ka \ll 1$, find the s-wave phase shift for the square well, and show that it has the expansion

$$k \cot\delta_0(k) = -\frac{1}{a_0} + \frac{1}{2} r_0\, k^2 + \cdots \tag{23.40}$$

Here a_0 is called the "scattering length," and r_0 the "effective range." Find the cross section using (23.40). Note that the scattering amplitude at $k = 0$ can be found directly from the s-wave Schrödinger's equation.

The behavior of the phase shifts as a function of k gives useful information about the general properties of the potential and the nature of the scattering process. Levinson's theorem relates the value of the phase shifts at $k = 0$ to the number of bound states for a given angular momentum ℓ:

Theorem 58 (Levinson's theorem)**.** *This curious theorem states that if the phase shifts are normalized such that $\delta_\ell(k) \to 0$ at $k \to \infty$, and if the phase shifts are then followed continuously into the origin where $k = 0$, the phase shift at $k = 0$ is given by a multiple of π,*

$$\delta_\ell(0) = N\pi \,, \tag{23.41}$$

where N is the number of bound states that the potential can support for angular momentum ℓ.

This theorem is illustrated for a square well for $\ell = 0$ and $\ell > 0$ in Fig. 1. Notice that for $\ell = 0$, if the potential does not quite bind, the phase shift never quite gets to $\pi/2$, whereas if the potential barely binds one state, the phase shift increases through $\pi/2$, reaching a value of π at the origin.

Discussion of scattering length.

For $\ell > 0$, if the potential doesn't quite bind, the phase shift can pass through $\pi/2$ with a negative slope then again with a positive slope, winding up at $\delta_\ell(0) = 0$, whereas if the potential barely binds one state, the phase shift again only increases through $\pi/2$. We will show below that we can associate a classical resonance with the case when the phase shift goes through $\pi/2$ with a *positive* slope.

We will see an illustration of this theorem in proton-neutron scattering at low energy in what follows next.

Example 45. Illustrate Levenson's theorem for the square well for low energy s-wave scattering by plotting the phase shift $\delta_0(k)$ *vs* k for the case when there is "almost" one bound state, and when there is one bound state.

23.4.1 Resonance and time delays

Here we introduce wave packets for the incident beam, and show that the asymptotic form of the wave function provides a description of the time dependence of the scattering process. We also discuss time delays and resonance phenomena here.

23.5 Proton-Neutron scattering

The potential between protons and neutrons is strongly dependent on the relative spin orientation of the particles. In fact, if the particles is in the singlet state they are unbound, but in the triplet state they can bind, and, in such a state, become the deuteron. The interaction is described by the potential,

$$V(|\mathbf{r}_1 - \mathbf{r}_2|, \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2) = V_0(|\mathbf{r}_1 - \mathbf{r}_2|)\, P_0(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2) + V_1(|\mathbf{r}_1 - \mathbf{r}_2|)\, P_1(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)\,, \tag{23.42}$$

where the singlet and triplet projection operators P_0 and P_1 are given by:

$$P_0(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2) = \qquad \chi_{00}(1,2)\, \chi^\dagger_{00}(1,2) \qquad = \frac{1 - \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2}{4}\,, \tag{23.43}$$

$$P_1(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2) = \sum_{M=-1}^{1} \chi_{1M}(1,2)\, \chi^\dagger_{1M}(1,2) = \frac{3 + \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2}{4}\,. \tag{23.44}$$

We follow the same method as described in the preceding section. We take particle 1 to be incident on particle 2 at rest, and change variables to relative and center of mass coordinates. Solutions to Schrödinger's equation can then be written as

$$\psi_{\mathbf{K},\mathbf{k}}(\mathbf{R}, \mathbf{r}) = e^{i\mathbf{K}\cdot\mathbf{R}}\, \psi_{\mathbf{k}}(\mathbf{r})\,, \tag{23.45}$$

$$\psi_{\mathbf{k}}(\mathbf{r}) = \sum_{SM} C_{SM}(\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2)\, \psi_{\mathbf{k},S}(\mathbf{r})\, \chi_{SM}(1,2)\,, \tag{23.46}$$

where $\psi_{\mathbf{k},S}(\mathbf{r})$ is the solution of

$$\left\{ -\frac{\hbar^2}{2\mu} \boldsymbol{\nabla}_r^2 + V_S(r) \right\} \psi_{\mathbf{k},S}(\mathbf{r}) = \frac{\hbar^2 k^2}{2\mu}\, \psi_{\mathbf{k},S}(\mathbf{r})\,. \tag{23.47}$$

with asymptotic conditions,

$$\psi_{\mathbf{k},S}(\mathbf{r}) \sim e^{i\mathbf{k}\cdot\mathbf{r}} + f_{k,S}(\hat{\mathbf{r}})\, \frac{e^{ikr}}{r}\,, \qquad \text{as } r \to \infty, \tag{23.48}$$

The scattering amplitudes $f_{k,S}(\hat{\mathbf{r}})$ are given in terms of the phase shifts $\delta_{\ell S}(k)$ by

$$f_{k,S}(\hat{\mathbf{r}}) = \frac{1}{k}\sum_{\ell=0}^{\infty}(2\ell+1)\,e^{i\delta_{\ell S}(k)}\sin\delta_{\ell S}(k)\,P_\ell(\cos\theta)\,, \tag{23.49}$$

which must be found by solving the radial part of Schrödinger's equation (23.47) for each S-dependent potential, exactly as in the spinless case discussed in the previous section.

The coefficients $C_{SM}(\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2)$ in (23.46) are to be fixed by the initial spin polarizations of the particles, $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$. Thus the incident and scattered waves are given by:

$$\psi_{\mathbf{k},\,\mathrm{inc}}(\mathbf{r}) = \sum_{SM} C_{SM}(\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2)\,\chi_{SM}(1,2)\,e^{i\mathbf{k}\cdot\mathbf{r}}\,, \tag{23.50}$$

$$\psi_{\mathbf{k},\,\mathrm{scat}}(\mathbf{r}) = \sum_{SM} C_{SM}(\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2)\,\chi_{SM}(1,2)\,f_{k,S}(\hat{\mathbf{r}})\,\frac{e^{ikr}}{r}\,. \tag{23.51}$$

We take the incident spinor to be the product of two spinors for each particle with polarization vectors $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$:

$$\chi_{\hat{\mathbf{p}}_1}(1) = \sum_{m_1} C_{m_1}(\hat{\mathbf{p}}_1)\,\chi_{m_1}(1) = \begin{pmatrix}\cos(\theta_1/2)\\ e^{i\phi_1}\sin(\theta_1/2)\end{pmatrix}\,,$$

$$\chi_{\hat{\mathbf{p}}_2}(2) = \sum_{m_2} C_{m_2}(\hat{\mathbf{p}}_2)\,\chi_{m_2}(2) = \begin{pmatrix}\cos(\theta_2/2)\\ e^{i\phi_2}\sin(\theta_2/2)\end{pmatrix}\,.$$

Then we have

$$\chi_{\hat{\mathbf{p}}_1}(1)\chi_{\hat{\mathbf{p}}_2}(2) = \sum_{SM} C_{SM}(\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2)\,\chi_{SM}(1,2)\,, \tag{23.52}$$

which we can invert to find the coefficients $C_{SM}(\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2)$:

$$C_{SM}(\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2) = \sum_{m_1,m_2}\langle\, m_1 m_2\,|\,SM\,\rangle\, C_{m_1}(\hat{\mathbf{p}}_1)\,C_{m_2}(\hat{\mathbf{p}}_2)\,. \tag{23.53}$$

The final spinor is given by:

$$\begin{aligned}
&f_k(\hat{\mathbf{r}},\hat{\mathbf{p}}_1',\hat{\mathbf{p}}_2',\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2)\,\chi_{\mathrm{scat}}(1,2)\\
&\quad= \sum_{SM} f_{k,S}(\hat{\mathbf{r}})\,C_{SM}(\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2)\,\chi_{SM}(1,2)\,,\\
&\quad= [\,f_{k,0}(\hat{\mathbf{r}})\,P_0(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2) + f_{k,1}(\hat{\mathbf{r}})\,P_1(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2)\,]\sum_{SM} C_{SM}(\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2)\,\chi_{SM}(1,2)\,,\\
&\quad= [\,f_{k,0}(\hat{\mathbf{r}})\,P_0(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2) + f_{k,1}(\hat{\mathbf{r}})\,P_1(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2)\,]\,\chi_{\hat{\mathbf{p}}_1}(1)\chi_{\hat{\mathbf{p}}_2}(2)\,.
\end{aligned}$$

The scattering amplitude $f_k(\hat{\mathbf{r}},\hat{\mathbf{p}}_1',\hat{\mathbf{p}}_2',\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2)$ is introduced here so that the scattered spinor $\chi_{\mathrm{scat}}(1,2)$ can be normalized to one. Thus if we take for the scattered spinor the form,

$$\chi_{\mathrm{scat}}(1,2) = \chi_{\hat{\mathbf{p}}_1'}(1)\chi_{\hat{\mathbf{p}}_2'}(2)\,, \tag{23.54}$$

the scattering amplitude $f_k(\hat{\mathbf{r}})$ is given by:

$$\begin{aligned}
&f_k(\hat{\mathbf{r}},\hat{\mathbf{p}}_1',\hat{\mathbf{p}}_2',\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2) =\\
&\qquad\chi^\dagger_{\hat{\mathbf{p}}_1'}(1)\chi^\dagger_{\hat{\mathbf{p}}_2'}(2)\,[\,f_{k,0}(\hat{\mathbf{r}})\,P_0(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2) + f_{k,1}(\hat{\mathbf{r}})\,P_1(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2)\,]\,\chi_{\hat{\mathbf{p}}_1}(1)\chi_{\hat{\mathbf{p}}_2}(2)\,,
\end{aligned} \tag{23.55}$$

and the differential scattering cross section is then given by:

$$\begin{aligned}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} &= |f_k(\hat{\mathbf{r}},\hat{\mathbf{p}}_1',\hat{\mathbf{p}}_2',\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2)|^2 \\ &= \mathrm{Tr}_1\mathrm{Tr}_2\Big\{\rho_{\hat{\mathbf{p}}_1'}(1)\rho_{\hat{\mathbf{p}}_2'}(2)[\,f_{k,0}^*(\hat{\mathbf{r}})\,P_0(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2)+f_{k,1}^*(\hat{\mathbf{r}})\,P_1(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2)\,] \\ &\qquad\times\rho_{\hat{\mathbf{p}}_1}(1)\rho_{\hat{\mathbf{p}}_2}(2)[\,f_{k,0}(\hat{\mathbf{r}})\,P_0(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2)+f_{k,1}(\hat{\mathbf{r}})\,P_1(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2)\,]\Big\} \qquad (23.56)\end{aligned}$$

where the density matricies ρ are given by:

$$\begin{aligned}\rho_{\hat{\mathbf{p}}_1}(1) &= \chi_{\hat{\mathbf{p}}_1}(1)\chi_{\hat{\mathbf{p}}_1}\dagger(1) = \frac{1}{2}(1+\hat{\mathbf{p}}_1\cdot\boldsymbol{\sigma}_1)\\ \rho_{\hat{\mathbf{p}}_2}(2) &= \chi_{\hat{\mathbf{p}}_2}(2)\chi_{\hat{\mathbf{p}}_2}\dagger(2) = \frac{1}{2}(1+\hat{\mathbf{p}}_2\cdot\boldsymbol{\sigma}_2)\\ \rho_{\hat{\mathbf{p}}_1'}(1) &= \chi_{\hat{\mathbf{p}}_1'}(1)\chi_{\hat{\mathbf{p}}_1'}\dagger(1) = \frac{1}{2}(1+\hat{\mathbf{p}}_1'\cdot\boldsymbol{\sigma}_1)\\ \rho_{\hat{\mathbf{p}}_2'}(2) &= \chi_{\hat{\mathbf{p}}_2'}(2)\chi_{\hat{\mathbf{p}}_2'}\dagger(2) = \frac{1}{2}(1+\hat{\mathbf{p}}_2'\cdot\boldsymbol{\sigma}_2)\end{aligned}$$

Eq. (23.56) is the formula we seek.

It is not clear that the final spinor is just a simple direct product of two spinors. This is the subject of the next example.

Example 46. Find the density matrix for the scattered wave,

$$\rho_{\mathrm{scatt}}(1,2) = \chi_{\mathrm{scat}}(1,2)\chi^\dagger_{\mathrm{scat}}(1,2)\,,$$

in terms of $f_{k,0}(\hat{\mathbf{r}})$, $f_{k,1}(\hat{\mathbf{r}})$, $\hat{\mathbf{p}}_1$, and $\hat{\mathbf{p}}_2$. Show that $\rho_{\mathrm{scatt}}(1,2)$ is given by the direct product of density matrices for the two particles:

$$\rho_{\mathrm{scatt}}(1,2) = \rho_{\hat{\mathbf{p}}_1'}(1)\rho_{\hat{\mathbf{p}}_2'}(2) = \frac{1}{2}(1+\hat{\mathbf{p}}_1'\cdot\boldsymbol{\sigma}_1)\,\frac{1}{2}(1+\hat{\mathbf{p}}_2'\cdot\boldsymbol{\sigma}_2)\,. \qquad (23.57)$$

Find $\hat{\mathbf{p}}_1'$ and $\hat{\mathbf{p}}_2'$. [This problem might be too difficult! I believe that the final density matrix is not just the simple product of two density matrices, and that there will be tensor correlations. Why don't the books discuss this?]

References

Part III

Appendices

Appendix A

Table of physical constants

speed of light in vacuum	c	$2.997\ 924\ 58 \times 10^{8}$ m/s
Planck constant	h	$6.626\ 069\ 3(11) \times 10^{-34}$ J-s
Planck constant, reduced	$\hbar \equiv h/(2\pi)$	$1.054\ 571\ 68(18) \times 10^{-34}$ J-s
		$6.582\ 119\ 15(56) \times 10^{-22}$ MeV-s
electron charge magnitude	e	$1.602\ 176\ 53(14) \times 10^{-19}$ C
		$4.803\ 204\ 41(41) \times 10^{-10}$ esu
conversion constant	$\hbar\, c$	197.326 968(17) eV-nm
		197.326 968(17) MeV-fm
electron mass	m_e	0.510 998 918(44) MeV/c^2
		$9.109\ 3826(16) \times 10^{-31}$ kg
proton mass	m_p	938.272 029(80) MeV/c^2
		$1.672\ 621\ 71(29) \times 10^{-27}$ kg
		1.007 276 466 88(13) u
		1836.152 672 61(85) m_e
Bohr magneton	$\mu_B = e\hbar/(2m_e c)$	$5.788\ 381\ 804(39) \times 10^{-11}$ MeV/T
		1.40 MHz/Gauss
nuclear magneton	$\mu_N = e\hbar/(2m_p c)$	$3.152\ 451\ 259(21) \times 10^{-14}$ MeV/T

Table A.1: Table of physical constants from the particle data group.

1. PHYSICAL CONSTANTS

Table 1.1. Reviewed 2005 by P.J. Mohr and B.N. Taylor (NIST). Based mainly on the "CODATA Recommended Values of the Fundamental Physical Constants: 2002" by P.J. Mohr and B.N. Taylor, Rev. Mod. Phys. **77**, 1 (2005). The last group of constants (beginning with the Fermi coupling constant) comes from the Particle Data Group. The figures in parentheses after the values give the 1-standard-deviation uncertainties in the last digits; the corresponding fractional uncertainties in parts per 10^9 (ppb) are given in the last column. This set of constants (aside from the last group) is recommended for international use by CODATA (the Committee on Data for Science and Technology). The full 2002 CODATA set of constants may be found at `http://physics.nist.gov/constants`

Quantity	Symbol, equation	Value	Uncertainty (ppb)
speed of light in vacuum	c	299 792 458 m s^{-1}	exact*
Planck constant	h	$6.626\ 0693(11)\times10^{-34}$ J s	170
Planck constant, reduced	$\hbar \equiv h/2\pi$	$1.054\ 571\ 68(18)\times10^{-34}$ J s	170
		$= 6.582\ 119\ 15(56)\times10^{-22}$ MeV s	85
electron charge magnitude	e	$1.602\ 176\ 53(14)\times10^{-19}$ C $= 4.803\ 204\ 41(41)\times10^{-10}$ esu	85, 85
conversion constant	$\hbar c$	197.326 968(17) MeV fm	85
conversion constant	$(\hbar c)^2$	0.389 379 323(67) GeV^2 mbarn	170
electron mass	m_e	$0.510\ 998\ 918(44)$ MeV/c^2 $= 9.109\ 3826(16)\times10^{-31}$ kg	86, 170
proton mass	m_p	938.272 029(80) MeV/c^2 $= 1.672\ 621\ 71(29)\times10^{-27}$ kg	86, 170
		$= 1.007\ 276\ 466\ 88(13)$ u $= 1836.152\ 672\ 61(85)\ m_e$	0.13, 0.46
deuteron mass	m_d	1875.612 82(16) MeV/c^2	86
unified atomic mass unit (u)	(mass ^{12}C atom)/12 = (1 g)/(N_A mol)	931.494 043(80) MeV/c^2 $= 1.660\ 538\ 86(28)\times10^{-27}$ kg	86, 170
permittivity of free space	$\epsilon_0 = 1/\mu_0 c^2$	$8.854\ 187\ 817\ \ldots\times10^{-12}$ F m^{-1}	exact
permeability of free space	μ_0	$4\pi \times 10^{-7}$ N A^{-2} $= 12.566\ 370\ 614\ \ldots\times10^{-7}$ N A^{-2}	exact
fine-structure constant	$\alpha = e^2/4\pi\epsilon_0\hbar c$	$7.297\ 352\ 568(24)\times10^{-3} = 1/137.035\ 999\ 11(46)^\dagger$	3.3, 3.3
classical electron radius	$r_e = e^2/4\pi\epsilon_0 m_e c^2$	$2.817\ 940\ 325(28)\times10^{-15}$ m	10
(e^- Compton wavelength)/2π	$\lambda\!\!\!^{-}_e = \hbar/m_e c = r_e\alpha^{-1}$	$3.861\ 592\ 678(26)\times10^{-13}$ m	6.7
Bohr radius ($m_{\rm nucleus} = \infty$)	$a_\infty = 4\pi\epsilon_0\hbar^2/m_e e^2 = r_e\alpha^{-2}$	$0.529\ 177\ 2108(18)\times10^{-10}$ m	3.3
wavelength of 1 eV/c particle	$hc/(1\text{ eV})$	$1.239\ 841\ 91(11)\times10^{-6}$ m	85
Rydberg energy	$hcR_\infty = m_e e^4/2(4\pi\epsilon_0)^2\hbar^2 = m_e c^2\alpha^2/2$	13.605 6923(12) eV	85
Thomson cross section	$\sigma_T = 8\pi r_e^2/3$	0.665 245 873(13) barn	20
Bohr magneton	$\mu_B = e\hbar/2m_e$	$5.788\ 381\ 804(39)\times10^{-11}$ MeV T^{-1}	6.7
nuclear magneton	$\mu_N = e\hbar/2m_p$	$3.152\ 451\ 259(21)\times10^{-14}$ MeV T^{-1}	6.7
electron cyclotron freq./field	$\omega^e_{\rm cycl}/B = e/m_e$	$1.758\ 820\ 12(15)\times10^{11}$ rad s^{-1} T^{-1}	86
proton cyclotron freq./field	$\omega^p_{\rm cycl}/B = e/m_p$	$9.578\ 833\ 76(82)\times10^{7}$ rad s^{-1} T^{-1}	86
gravitational constant‡	G_N	$6.6742(10)\times10^{-11}$ m^3 kg^{-1} s^{-2}	1.5×10^5
		$= 6.7087(10)\times10^{-39}\ \hbar c$ $(\text{GeV}/c^2)^{-2}$	1.5×10^5
standard gravitational accel.	g_n	9.806 65 m s^{-2}	exact
Avogadro constant	N_A	$6.022\ 1415(10)\times10^{23}$ mol^{-1}	170
Boltzmann constant	k	$1.380\ 6505(24)\times10^{-23}$ J K^{-1}	1800
		$= 8.617\ 343(15)\times10^{-5}$ eV K^{-1}	1800
molar volume, ideal gas at STP	$N_A k(273.15\text{ K})/(101\ 325\text{ Pa})$	$22.413\ 996(39)\times10^{-3}$ m^3 mol^{-1}	1700
Wien displacement law constant	$b = \lambda_{\rm max}T$	$2.897\ 7685(51)\times10^{-3}$ m K	1700
Stefan-Boltzmann constant	$\sigma = \pi^2 k^4/60\hbar^3 c^2$	$5.670\ 400(40)\times10^{-8}$ W m^{-2} K^{-4}	7000
Fermi coupling constant**	$G_F/(\hbar c)^3$	$1.166\ 37(1)\times10^{-5}$ GeV^{-2}	9000
weak-mixing angle	$\sin^2\widehat{\theta}(M_Z)$ $(\overline{\rm MS})$	$0.23122(15)^{\dagger\dagger}$	6.5×10^5
$W^\pm$ boson mass	m_W	80.403(29) GeV/c^2	3.6×10^5
Z^0 boson mass	m_Z	91.1876(21) GeV/c^2	2.3×10^4
strong coupling constant	$\alpha_s(m_Z)$	0.1176(20)	1.7×10^7

$\pi = 3.141\ 592\ 653\ 589\ 793\ 238$ e = 2.718 281 828 459 045 235 $\gamma = 0.577\ 215\ 664\ 901\ 532\ 861$

1 in $\equiv 0.0254$ m 1 G $\equiv 10^{-4}$ T 1 eV $= 1.602\ 176\ 53(14)\times10^{-19}$ J kT at 300 K $= [38.681\ 684(68)]^{-1}$ eV

1 Å $\equiv 0.1$ nm 1 dyne $\equiv 10^{-5}$ N 1 eV/$c^2 = 1.782\ 661\ 81(15)\times10^{-36}$ kg 0 °C $\equiv 273.15$ K

1 barn $\equiv 10^{-28}$ m^2 1 erg $\equiv 10^{-7}$ J $2.997\ 924\ 58\times10^9$ esu = 1 C 1 atmosphere $\equiv 760$ Torr $\equiv 101\ 325$ Pa

* The meter is the length of the path traveled by light in vacuum during a time interval of 1/299 792 458 of a second.
† At $Q^2 = 0$. At $Q^2 \approx m_W^2$ the value is $\sim 1/128$.
‡ Absolute lab measurements of G_N have been made only on scales of about 1 cm to 1 m.
** See the discussion in Sec. 10, "Electroweak model and constraints on new physics."
†† The corresponding $\sin^2\theta$ for the effective angle is 0.23152(14).

Table A.2: Table of physical constants from the particle data group.

Appendix B

Operator Relations

B.1 Commutator identities

The commutator and anti-commutator of two operators are written thus:

$$[\,A,B\,] = AB - BA\,,$$
$$\{\,A,B\,\} = AB + BA\,.$$

Sometimes we use the notation $[\,A,B\,]_{\mp}$ for commutators $(-)$ and anti-commutators $(+)$. It is easy to verify the following elementary commutator relations:

$$\begin{gathered}
[\,A,\alpha B+\beta C\,] = \alpha\,[\,A,B\,] + \beta\,[\,A,C\,] \\
[\,A,BC\,] = [\,A,B\,]\,C + B\,[\,A,C\,] \\
[\,AB,C\,] = A\,[\,B,C\,] + [\,A,C\,]\,B \\
[\,AB,CD\,] = A\,[\,B,C\,]\,D + [\,A,C\,]\,BD + CA\,[\,B,D\,] + C\,[\,A,D\,]\,B \\
[\,A,[\,B,C\,]\,] + [\,B,[\,C,A\,]\,] + [\,C,[\,A,B\,]\,] = 0 \qquad \text{(Jacobi's identity)} \\
[\,A,B^n\,] = nB^{n-1}\,[\,A,B\,] \\
[\,A^n,B\,] = nA^{n-1}\,[\,A,B\,]
\end{gathered}$$

For the special case of Q and P, obeying $[Q,P] = i\hbar$, we find:

$$[\,P,F(Q)\,]/i\hbar = -\frac{dF(Q)}{dQ} \tag{B.1}$$

$$[\,Q,F(P)\,]/i\hbar = +\frac{dF(P)}{dP} \tag{B.2}$$

If $[\,A,A^\dagger\,] = c$ where c is a number and $A\,|\,0\,\rangle = 0$, a useful identity is:

$$[\,(\,A\,)^n,(\,A^\dagger\,)^m\,]\,|\,0\,\rangle = \frac{m!}{(m-n)!}\,c^n\,(\,A^\dagger\,)^{m-n}\,|\,0\,\rangle\,, \qquad \text{for } m \geq n, \tag{B.3}$$

from which we also find:

$$\langle\,0\,|\,[\,(\,A\,)^n,(\,A^\dagger\,)^m\,]\,|\,0\,\rangle = \delta_{n,m}\,n!\,c^n\,. \tag{B.4}$$

B.2 Operator functions

Functions f of operators are defined by their power series expansions:

$$f(A) = \sum_{n=0}^{\infty} C_n \, A^n \, , \tag{B.5}$$

where C_n is a number. A function f of a set of *commuting* operators A_i, $i = 1, 2, \ldots, N$ is defined by a multiple power series expansion:

$$f(A_1, A_2, \ldots, A_N) = \sum_{n_1, n_2, \ldots, n_N = 0}^{\infty} C_{n_1, n_2, \ldots, n_N} \, A_1^{n_1} \, A_2^{n_2} \cdots A_N^{n_N} \, . \tag{B.6}$$

Definition 42 (Homogeneous functions)**.** Let $f(\, A_1, A_2, \ldots, A_N \,)$ be a smooth function of N commuting operators A_i, $i = 1, 2, \ldots, N$. It is a homogeneous function of degree n if it obeys the relation:

$$f(\, \lambda A_1, \lambda A_2, \ldots, \lambda A_N \,) = \lambda^n \, f(\, A_1, A_2, \ldots, A_N \,) \, . \tag{B.7}$$

In terms of its power series expansion, the coefficient $C_{n_1, n_2, \ldots, n_N}$ for a homogeneous function vanishes unless $n_1 + n_2 + \cdots + n_N = n$.

Example 47. The exponential operator $B = e^A$ is defined by the power series:

$$B = e^A = 1 + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots = \lim_{N \to \infty} \left(1 + \frac{A}{N} \right)^N \, . \tag{B.8}$$

Determinants and traces of operators are defined in terms of their eigenvalues. For example, for B defined above,

$$\det\{B\} = \prod_i b_i = e^{(\sum_i a_i)} = e^{\mathrm{Tr}\,\{A\}} \, . \tag{B.9}$$

We note the determinant and trace properties. For and operators A and B, we have

$$\det\{AB\} = \det\{A\} \det\{B\} \, , \tag{B.10}$$

$$\mathrm{Tr}\,\{AB\} = \mathrm{Tr}\,\{BA\} \, , \tag{B.11}$$

$$\mathrm{Tr}\,\{A + B\} = \mathrm{Tr}\,\{A\} + \mathrm{Tr}\,\{B\} \, . \tag{B.12}$$

If U is unitary and H hermitian, then

$$U = e^{iH} = \frac{1 + i \tan(H/2)}{1 - i \tan(H/2)} \, . \tag{B.13}$$

B.3 Operator theorems

Theorem 59 (Baker-Campbell-Hausdorff)**.** *In general, we have the expansion:*

$$e^A B e^{-A} = B + [A, B] + \frac{1}{2!}[A, [A, B]] + \frac{1}{3!}[A, [A, [A, B]]] + \cdots \tag{B.14}$$

Proof. We follow the proof given in Merzbacher [1][page 167]. Let $f(\lambda)$ be given by

$$f(\lambda) = e^{\lambda A} B e^{-\lambda A} = f(0) + \frac{f'(0)}{1!}\lambda + \frac{f''(0)}{2!}\lambda^2 + \frac{f'''(0)}{3!}\lambda^3 + \cdots$$

Then we find:

$$f(0) = B$$
$$f'(0) = [A, B]$$
$$f''(0) = [A, [A, B]] \,, \qquad \text{etc.}$$

Setting $\lambda = 1$ proves the theorem. □

Theorem 60 (exponent rule)**.** *This formula states that:*

$$\exp\{A\}\exp\{B\} = \exp\left\{A + B + \frac{1}{2}\,[A, B] + \frac{1}{12}\,\left(\,[A, [A, B]] - [B, [B, A]]\,\right) + \cdots\right\} . \tag{B.15}$$

If A and B both commute with their commutator,

$$[A, [A, B]] = [B, [A, B]] = 0 \,,$$

then the formula states that:

$$e^{A+B} = e^A e^B e^{-\frac{1}{2}\,[A,B]} = e^B e^A e^{+\frac{1}{2}\,[A,B]} \,. \tag{B.16}$$

Proof. We can use the proof given in Merzbacher[1][Exercise 8.18, p. 167]) for the case when $[A, [A, B]] = [B, [A, B]] = 0$. For this case, we let

$$f(\lambda) = e^{\lambda A} e^{\lambda B} e^{-\lambda(A+B)} \,.$$

Differenting $f(\lambda)$ with respect to λ gives:

$$\frac{\mathrm{d}f(\lambda)}{\mathrm{d}\lambda} = \lambda[A, B]f(\lambda) \,.$$

Solution of this differential equation gives (B.16). □

Theorem 61 (Euler's theorem on homogeneous functions)**.** *Let $f(A_1^\dagger, A_2^\dagger, \ldots, A_m^\dagger)$ be a homogeneous function of degree n of m commuting creation operators which obey the relation: $[\,A_i\,, A_j^\dagger\,] = \delta_{i,j}$. Then*

$$[\,N, f(A_1^\dagger, A_2^\dagger, \ldots, A_m^\dagger)\,] = n\, f(A_1^\dagger, A_2^\dagger, \ldots, A_m^\dagger) \,, \tag{B.17}$$

where

$$N = \sum_{i=1}^{m} A_i^\dagger\, A_i \,.$$

Proof. The power series expansion of f is:

$$f(A_1^\dagger, A_2^\dagger, \ldots, A_m^\dagger) = \sum_{n_1, n_2, \ldots, n_m = 0}^{\infty} C_{n_1, n_2, \ldots, n_m}\, \delta_{n, n_1+n_2+\cdots+n_m}\, {A_1^\dagger}^{n_1}\, {A_2^\dagger}^{n_2} \cdots {A_m^\dagger}^{n_m} \,. \tag{B.18}$$

So since

$$[\,A_i^\dagger\, A_i\,, {A_1^\dagger}^{n_1}\, {A_2^\dagger}^{n_2} \cdots {A_m^\dagger}^{n_m}\,] = n_i\, {A_1^\dagger}^{n_1}\, {A_2^\dagger}^{n_2} \cdots {A_m^\dagger}^{n_m} \,, \tag{B.19}$$

and since the function is homogeneous of degree n, $\sum_{i=1}^m n_i = n$, which proves the theorem. The *converse* of this theorem is also true. That is if f obeys Eq. (B.17), then it is a homogeneous function of degree n of the m operators $A_i^\dagger$. That is $f(A_1^\dagger, A_2^\dagger, \ldots, A_m^\dagger)\,|\,0\,\rangle$ is an eigenvector of N with eigenvalue n:

$$N\, f(A_1^\dagger, A_2^\dagger, \ldots, A_m^\dagger)\,|\,0\,\rangle = n\, f(A_1^\dagger, A_2^\dagger, \ldots, A_m^\dagger)\,|\,0\,\rangle \,. \tag{B.20}$$

[Note: Euler's theorem is usually stated in terms of a function $f(x_1, x_2, \ldots, x_m)$ of real variables x_i where N is given by the differential operator $N \mapsto \sum_{i=1}^m x_i\, \partial/\partial x_i$. This is an *operator* version of the same theorem.] □

References

[1] E. Merzbacher, *Quantum Mechanics* (John Wiley & Sons, New York, NY, 1970), second edition.

Appendix C

Binomial coefficients

In this appendix we review some relations between binomial coefficients which we use in Chapter **??**. Reference material for binomial coefficients can be found on the web (Wikipedia), and Appendix 1 in Edmonds [1].

The binomial coefficient is defined to be the coefficient of powers of x and y in the expansion of $(x+y)^n$:

$$(x+y)^n = \sum_m \binom{n}{m} x^{n-m}\, y^m = \sum_m \binom{n}{m} x^m\, y^{n-m}\,. \tag{C.1}$$

For a positive integer $n \geq 0$, the binomial coefficient is given by:

$$\binom{n}{m} = \begin{cases} \dfrac{n!}{(n-m)!\,m!}\,, & \text{for } n \geq 0 \text{ and } m \geq 0. \\ 0\,, & \text{for } n \geq 0 \text{ and } m < 0 \text{ or } m > n. \end{cases} \tag{C.2}$$

So for $n \geq 0$, the sum in Eq. (C.1) runs from $m = 0$ to $m = n$. For negative integers $n < 0$, the binomial coefficient is defined by:

$$\binom{n}{m} = (-)^m \binom{m-n-1}{m} = (-)^m \frac{(m-n-1)!}{(-n-1)!\,m!}\,, \qquad m \geq 0\,. \tag{C.3}$$

From (C.1),

$$\binom{n}{m} = \binom{n}{n-m}\,. \tag{C.4}$$

A recursion formula is Pascal's rule:

$$\binom{n}{m} + \binom{n}{m+1} = \binom{n+1}{m+1}\,, \tag{C.5}$$

which can be proved for $n \geq 0$ using the definiton (C.2) by manipulation of the factorials. By considering the identity $(x+y)^n\,(x+y)^m = (x+y)^{n+m}$, we find Vandermonde's identity:

$$\sum_k \binom{n}{k} \binom{m}{l-k} = \binom{n+m}{l}\,, \tag{C.6}$$

which, for $n > 0$ and $m > 0$, gives the relation:

$$\sum_k \frac{1}{(n-k)!\,k!\,(m-l+k)!\,(l-k)!} = \frac{(n+m)!}{n!\,m!\,(n+m-l)!\,l!}\,. \tag{C.7}$$

For $n < 0$ and $m < 0$, Vendermonde's identity becomes:

$$\sum_k \binom{k-n-1}{k}\binom{l-k-m-1}{l-k} = \binom{l-n-m-1}{l}\,, \tag{C.8}$$

which gives the relation:

$$\sum_k \frac{(k-n-1)!\,(l-k-m-1)!}{k!\,(l-k)!} = \frac{(-n-1)!\,(-m-1)!\,(l-n-m-1)!}{(-n-m-1)!\,l!}\,. \tag{C.9}$$

Setting $k = s + c$, Eq. (C.9) becomes:

$$\sum_s \frac{(c-n-1+s)!\,(l-m-1-c-s)!}{(c+s)!\,(l-c-s)!} = \frac{(-n-1)!\,(-m-1)!\,(l-n-m-1)!}{(-n-m-1)!\,l!}\,. \tag{C.10}$$

Furthermore, setting $c - n - 1 = a$, $l - m - 1 - c = b$, and $l - c = d$ so that (C.10) becomes:

$$\sum_s \frac{(a+s)!\,(b-s)!}{(c+s)!\,(d-s)!} = \frac{(a-c)!\,(b-d)!\,(a+b+1)!}{(a+b-c-d+1)!\,(d+c)!}\,. \tag{C.11}$$

Setting $a = c = 0$ gives the relation:

$$\sum_s \frac{(b-s)!}{(d-s)!} = \frac{(b-d)!\,(b+1)!}{(b-d+1)!\,d!}\,, \tag{C.12}$$

whereas setting $b = d = 0$ gives:

$$\sum_s \frac{(a+s)!}{(c+s)!} = \frac{(a-c)!\,(a+1)!}{(a-c+1)!\,c!}\,. \tag{C.13}$$

For $-n > m \geq 0$, Vandermonde's identity becomes:

$$\sum_k (-)^k \binom{k-n-1}{k}\binom{m}{l-k} = (-)^l \binom{l-n-m-1}{l}\,, \tag{C.14}$$

which gives the relation:

$$\sum_k (-)^k \frac{(k-n-1)!}{k!\,(m-l+k)!\,(l-k)!} = (-)^l \frac{(-n-1)!\,(l-n-m-1)!}{m!\,(-n-m-1)!\,l!}\,. \tag{C.15}$$

Setting $-n - 1 = a$, $m - l = b$, and $l = c$, Eq. (C.15) becomes:

$$\sum_k (-)^k \frac{(a+k)!}{k!\,(b+k)!\,(c-k)!} = (-)^c \frac{a!\,(a-b)!}{c!\,(b+c)!\,(a-b-c)!}\,, \tag{C.16}$$

for $b \geq 0$ and $a - b \geq c \geq 0$.

References

[1] A. R. Edmonds, *Angular momentum in quantum mechanics* (Princeton University Press, Princeton, NJ, 1996), fourth printing with corrections, second edition.

ANNOTATION: This printing corrected several major errors in Chapter 4 in earlier printings.

Appendix D

Fourier transforms

D.1 Finite Fourier transforms

In this section, we follow the convensions of Numerical Recipes [1, p. ??]. For any complex function f_n of an integer n which is periodic in n with period N so that $f_n = f_{n+N}$, we can define a finite Fourier transform $\tilde{f}_k$ by the definitions:

$$f_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \tilde{f}_k \, e^{+2\pi i \, nk/N} \,, \tag{D.1}$$

$$\tilde{f}_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f_n \, e^{-2\pi i \, nk/N} \,. \tag{D.2}$$

The inverse relation (D.2) can be obtained from (D.1) by first noting that:

$$\sum_{n=0}^{N-1} x^n = 1 + x + x^2 + \cdots + x^{N-1} = \frac{1 - x^N}{1 - x} \,. \tag{D.3}$$

So setting $x = \exp\{2\pi i \, (k - k')/N\}$, we find

$$\sum_{n=0}^{N-1} e^{2\pi i \, (k-k')/N} = \frac{1 - e^{2\pi i \, (k-k')}}{1 - e^{2\pi i \, (k-k')/N}} = N \, \delta_{k,k'} \,, \tag{D.4}$$

from which (D.2) follows. The same trick can be used to derive (D.1) from (D.2). The normalization of $\tilde{f}_k$ can be defined in other ways (see ref. [1]).

Since $\tilde{f}_k$ is also periodic in k with period N, it is often useful to change the range of k by defining k' by:

$$k' = \begin{cases} k \,, & \text{for } 0 \le k \le [N/2], \\ k - N \,, & \text{for } [N/2] < k \le N - 1. \end{cases} \tag{D.5}$$

So the range of k' is from $-[N/2] + 1 \le k \le [N/2]$. The reverse relation is:

$$k = \begin{cases} k' \,, & \text{for } 0 \le k' \le [N/2], \\ k' + N \,, & \text{for } 0 > k' \ge -[N/2] + 1. \end{cases} \tag{D.6}$$

Here $[N/2]$ means the largest integer of $N/2$. For fast Fourier transform routines, we must select N to be a power of 2 so that N is always even. Then since

$$e^{+2\pi i \, nk/N} = e^{+2\pi i \, nk'/N} \,,$$

the finite Fourier transform pair can also be written as:

$$\begin{aligned} f_n &= \frac{1}{\sqrt{N}} \sum_{k'=-[N/2]+1}^{[N/2]} \tilde{f}_{k'}\, e^{+2\pi i\, nk'/N}\ , \\ \tilde{f}_{k'} &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f_n\, e^{-2\pi i\, nk'/N}\ . \end{aligned} \tag{D.7}$$

We often drop the prime notation and just put $k' \mapsto k$.

D.2 Finite sine and cosine transforms

It is sometimes useful to have finite sine and cosine transforms. These can be generated from the general transforms above.

References

[1] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN: The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 1992).

Appendix E

Classical mechanics

In Section E.1 in this chapter, we review the Lagrangian, Hamiltonian, and Poisson bracket formulations of classical mechanics. In Sections E.2, we review differential geometry and the symplectic formulation of classical mechanics.

E.1 Lagrangian and Hamiltonian dynamics

In this section, we use the notation of Goldstein [1] and Tabor [2][p. 48].

We suppose we can describe the dynamics of a system by n generalized coordinates $q \equiv (q_1, q_2, \ldots, q_n)$ and an action $S[q]$ defined by:

$$S[q] = \int \mathrm{d}t\, L(q, \dot{q})\,, \tag{E.1}$$

where $L(q, \dot{q})$ is the Lagrangian, which is at most quadratic in $\dot{q}_i$. Variation of the action leads to Lagrange's equations of motion:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial L(q,\dot{q})}{\partial \dot{q}_i}\Big] - \frac{\partial L(q,\dot{q})}{\partial q_i} = 0\,, \qquad i = 1, 2, \ldots, n. \tag{E.2}$$

Canonical momenta p_i are defined by:

$$p_i = \frac{\partial L(q,\dot{q})}{\partial \dot{q}_i}\,, \qquad i = 1, 2, \ldots, n, \tag{E.3}$$

which, if the determinant of the matrix of *second* derivatives of the Lagrangian is nonsingular, can be solved in the inverse way for $p_i = p_i(q, \dot{q})$. A Hamiltonian can then be defined by the Legendre transformation [2][For details, see p. 79.]:

$$H(q,p) = \sum_i p_i \dot{q}_i - L(q,\dot{q})\,. \tag{E.4}$$

Hamilton's equations of motion are then given by:

$$\dot{q}_i = +\frac{\partial H(q,p)}{\partial p_i}\,, \qquad \dot{p}_i = -\frac{\partial H(q,p)}{\partial q_i}\,. \tag{E.5}$$

Poisson brackets are defined by:

$$\{\, A(q,p), B(q,p) \,\} = \sum_i \Big\{ \frac{\partial A(q,p)}{\partial q_i}\frac{\partial B(q,p)}{\partial p_i} - \frac{\partial B(q,p)}{\partial q_i}\frac{\partial A(q,p)}{\partial p_i} \Big\}\,. \tag{E.6}$$

In particular,

$$\{\, q_i, p_j \,\} = \delta_{ij}\,, \qquad \{\, q_i, q_j \,\} = 0\,, \qquad \{\, p_i, p_j \,\} = 0\,. \tag{E.7}$$

The time derivative of any function of q, p is given by:

$$\begin{aligned}\frac{\mathrm{d}A(q,p,t)}{\mathrm{d}t} &= \sum_i \left\{ \frac{\partial A}{\partial q_i}\,\dot{q}_i + \frac{\partial A}{\partial p_i}\,\dot{p}_i \right\} + \frac{\partial A}{\partial t} \\ &= \{\, A, H \,\} + \frac{\partial A}{\partial t}\,.\end{aligned} \tag{E.8}$$

Using Poisson brackets, Hamilton's equations can be written as:

$$\dot{q}_i = \{\, q_i, H(q,p) \,\}\,, \qquad \dot{p}_i = \{\, p_i, H(q,p) \,\}\,. \tag{E.9}$$

Long ago it was recognized that one could introduce a matrix, or **symplectic**,[1] form of Hamilton's equations by defining $2n$ components of a contra-variant vector x with the definition:

$$x = (q,p) = (q_1, q_2, \ldots, q_n; p_1, p_2, \ldots, p_n)\,. \tag{E.10}$$

For the symplectic coordinate x^μ, we use Greek indices which run from $\mu = 1, 2, \ldots, 2n$. Using these coordinates, Hamilton's equations can be written in component form as:

$$\dot{x}^\mu = f^{\mu\nu}\,\partial_\nu H(x)\,, \qquad \partial_\nu \equiv \frac{\partial}{\partial x^\nu}\,, \tag{E.11}$$

and where $f^{\mu\nu}$ are components of a antisymmetric $2n \times 2n$ matrix of the block form:

$$f^{\mu\nu} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\,. \tag{E.12}$$

The equations of motion (E.11) for the symplectic variables x^μ are a set of first order coupled equations, intertwined by the symplectic matrix $f^{\mu\nu}$. Using this notation, the Poisson bracket, defined in Eq. (E.6), is written in contra-variant coordinates x^μ as:

$$\{\, A(x), B(x) \,\} = (\,\partial_\mu A(x)\,)\, f^{\mu\nu}\, (\,\partial_\nu B(x)\,)\,. \tag{E.13}$$

In particular, Eqs. (E.7) are written as:

$$\{\, x^\mu, x^\nu \,\} = f^{\mu\nu}\,. \tag{E.14}$$

Now let us consider a general mapping from one set of x coordinates to another set X, given by: $X^\mu = X^\mu(x)$. Then the differentials and partial derivatives transform as:

$$\mathrm{d}X^\mu = \frac{\partial X^\mu}{\partial x^\nu}\,\mathrm{d}x^\nu\,, \qquad \frac{\partial}{\partial X^\mu} = \frac{\partial x^\nu}{\partial X^\mu}\,\frac{\partial}{\partial x^\nu}\,. \tag{E.15}$$

The Poisson brackets in Eq. (E.20) become:

$$\begin{aligned}\{\, A(X), B(X) \,\} &= \frac{\partial A(X)}{\partial x^\mu}\, f^{\mu\nu}\, \frac{\partial B(X)}{\partial x^\nu} \\ &= \frac{\partial A(X)}{\partial X^{\mu'}}\,\frac{\partial X^{\mu'}}{\partial x^\mu}\, f^{\mu\nu}\, \frac{\partial B(X)}{\partial X^{\nu'}}\,\frac{\partial X^{\nu'}}{\partial x^\nu} \\ &= \frac{\partial A(X)}{\partial X^{\mu'}}\, F^{\mu'\nu'}(X)\, \frac{\partial B(X)}{\partial X^{\nu'}}\,,\end{aligned} \tag{E.16}$$

where

$$F^{\mu'\nu'}(X) = \frac{\partial X^{\mu'}}{\partial x^\mu}\,\frac{\partial X^{\nu'}}{\partial x^\nu}\, f^{\mu\nu}\,. \tag{E.17}$$

[1] According to Goldstein [1][p. 343], the term "symplectic" comes from the Greek word for "intertwined." The word was apparently introduced by H. Weyl in his 1939 book on *Classical Groups*.

Now if the new matrix $F^{\mu'\nu'}(X) = f^{\mu'\nu'}$, *i.e.* it is again a constant matrix of the block form given in (E.12), then the transformation $X^\mu = X^\mu(x)$ is called **canonical** since we can identify new position Q and momentum coordinates P by setting $X = (Q, P)$. These new coordinates satisfy the same equations of motion and fundamental Poisson brackets,

$$\dot{X}^\mu = f^{\mu\nu}\,\frac{\partial H(X)}{\partial X^\nu}\,, \qquad \{\, X^\mu, X^\nu \,\} = f^{\mu\nu}\,. \tag{E.18}$$

Poisson brackets, as defined in Eq. (E.6) or Eq. (E.20) satisfy Jacobi's identity. Let $A(x)$, $B(x)$, and $C(x)$ be functions of x. Then Jacobi's identity is:

$$\{\, A, \{\, B, C \,\}\,\} + \{\, B, \{\, C, A \,\}\,\} + \{\, C, \{\, A, B \,\}\,\} = 0\,. \tag{E.19}$$

However, it is not necessary that $f^{\mu\nu}(x)$ be *independent* of x in order that Poisson brackets satisfy Jacobi's identity. Let us define in general:

$$\{\, A(x), B(x) \,\} = (\,\partial_\mu A(x)\,)\, f^{\mu\nu}(x)\,(\,\partial_\nu B(x)\,)\,, \tag{E.20}$$

with $f^{\mu\nu}(x)$ an antisymmetric non-singular matrix. We write the inverse matrix $f_{\mu\nu}(x)$ so that

$$f_{\mu\sigma}(x)\, f^{\sigma\nu}(x) = \delta^\nu_\mu\,. \tag{E.21}$$

We use $f_{\mu\nu}(x)$ to define covariant coordinates. We write:

$$\partial^\mu = f^{\mu\nu}(x)\,\partial_\nu\,, \qquad \mathrm{d}x_\mu = \mathrm{d}x^\nu\, f_{\nu\mu}(x)\,. \tag{E.22}$$

Note the differences in the *order* of the indices for raising and lowering indices for differentials and partial derivatives.[2] Then we can prove the following theorem:

Theorem 62 (Jacobi's identity)**.** *Poisson brackets, defined by:*

$$\{\, A(x), B(x) \,\} = (\,\partial_\mu A(x)\,)\, f^{\mu\nu}(x)\,(\,\partial_\nu B(x)\,)\,, \tag{E.23}$$

satisfy Jacobi's identity:

$$\{\, A(x), \{\, B(x), C(x) \,\}\,\} + \{\, B(x), \{\, C(x), A(x) \,\}\,\} + \{\, C(x), \{\, A(x), B(x) \,\}\,\} = 0\,, \tag{E.24}$$

if $f^{\mu\nu}(x)$ satisfies:

$$\partial_\mu\, f_{\nu\lambda}(x) + \partial_\nu\, f_{\lambda\mu}(x) + \partial_\lambda\, f_{\mu\nu}(x) = 0\,, \tag{E.25}$$

for three arbitrary functions $A(x)$, $B(x)$, and $C(x)$. We will show later that Eq. (E.25) implies that $\tilde{\mathrm{d}}\tilde{f} = 0$, which is Bianchi's identity for the symplectic two-form.

Proof. We first calculate (we assume here that A, B, C, and f all depend on x):

$$\begin{aligned}
&\{\, A, \{\, B, C \,\}\,\} + \{\, B, \{\, C, A \,\}\,\} + \{\, C, \{\, A, B \,\}\,\} \\
&\qquad = (\,\partial_\mu A\,)\,(\,\partial_\nu B\,)\,(\,\partial_\lambda C\,)\,\{\, f^{\mu\gamma}(\,\partial_\gamma f^{\nu\lambda}\,) + f^{\nu\gamma}(\,\partial_\gamma f^{\lambda\mu}\,) + f^{\lambda\gamma}(\,\partial_\gamma f^{\mu\nu}\,)\,\} \\
&\qquad\qquad = (\,\partial_\mu A\,)\,(\,\partial_\nu B\,)\,(\,\partial_\lambda C\,)\,\{\, (\,\partial^\mu f^{\nu\lambda}\,) + (\,\partial^\nu f^{\lambda\mu}\,) + (\,\partial^\lambda f^{\mu\nu}\,)\,\}\,.
\end{aligned} \tag{E.26}$$

But now we note that since $f^{\nu\lambda} f_{\lambda\gamma} = \delta^\nu_\gamma$, differentiating this expression with respect to ∂_μ, we find:

$$(\,\partial_\mu f^{\nu\lambda}\,)\, f_{\lambda\gamma} + f^{\nu\lambda}\,(\,\partial_\mu f_{\lambda\gamma}\,) = 0\,. \tag{E.27}$$

[2] We must be careful here because unlike the metric used in special and general relativity, $f^{\mu\nu}(x)$ is *anti*-symmetric.

Inverting this expression, and interchanging indices, we find:

$$(\,\partial_\mu f^{\nu\lambda}\,) = -f^{\nu\nu'}\,(\,\partial_\mu f_{\nu'\lambda'}\,)\,f^{\lambda'\lambda} = f^{\nu\nu'}\,f^{\lambda\lambda'}\,(\,\partial_\mu f_{\nu'\lambda'}\,)\,,$$
$$(\,\partial^\mu f^{\nu\lambda}\,) = f^{\mu\mu'}\,(\,\partial_{\mu'} f_{\nu\lambda}\,) = f^{\mu\mu'}\,f^{\nu\nu'}\,f^{\lambda\lambda'}\,(\,\partial_\mu f_{\nu'\lambda'}\,)\,.$$

Using this expression in the last line of Eq. (E.26), we find:

$$\{\,A,\{\,B,C\,\}\,\}+\{\,B,\{\,C,A\,\}\,\}+\{\,C,\{\,A,B\,\}\,\}$$
$$=(\,\partial^\mu A\,)\,(\,\partial^\nu B\,)\,(\,\partial^\lambda C\,)\,\{\,(\,\partial_\mu f_{\nu\lambda}\,)+(\,\partial_\nu f_{\lambda\mu}\,)+(\,\partial_\lambda f_{\mu\nu}\,)\,\}\,. \quad \text{(E.28)}$$

So if $f_{\mu\nu}(x)$ satisfies Eq. (E.25), Bianchi's identity:

$$\partial_\mu\, f_{\nu\lambda}(x)+\partial_\nu\, f_{\lambda\mu}(x)+\partial_\lambda\, f_{\mu\nu}(x)=0\,,$$

the functions $A(x)$, $B(x)$, and $C(x)$ satisfy Jacobi's identity, which was what we wanted to prove. □

The crucial point here is that $f^{\mu\nu}(x)$ is anti-symmetric, invertible, and satisfies Bianchi's identity. Then our definition of Poisson brackets in Eq. (E.23) satisfies Jacobi's identity. We can understand the origin of $f^{\mu\nu}(x)$ if we consider a Lagrangian of the form:

$$L(x,\dot{x})=\pi_\nu(x)\,\dot{x}^\nu-H(x)\,, \quad \text{(E.29)}$$

Then

$$\frac{\partial L}{\partial \dot{x}^\nu}=\pi_\nu(x)\,,\qquad \frac{\partial L}{\partial x^\nu}=(\,\partial_\nu\pi_\mu(x)\,)\,\dot{x}^\mu-\partial_\nu H(x)\,, \quad \text{(E.30)}$$

so that:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\,\frac{\partial L}{\partial\dot{x}^\nu}\,\right]=(\partial_\mu\,\pi_\nu(x)\,)\,\dot{x}^\mu\,, \quad \text{(E.31)}$$

and Lagrange's equation becomes:

$$f_{\nu\mu}(x)\,\dot{x}^\mu=\partial_\nu H(x)\,,\qquad \text{where}\qquad f_{\mu\nu}(x)=\partial_\mu\pi_\nu(x)-\partial_\nu\pi_\mu(x)\,. \quad \text{(E.32)}$$

From the definition of $f_{\mu\nu}(x)$ in Eq. (E.32) in terms of derivatives of $\pi_\nu(x)$, we see that it is antisymmetric. Also, in order to solve Lagrange's equations of motion, $f_{\nu\mu}(x)$ must be invertible. Satisfying Bianchi's identity is a further condition that must be imposed for the set (x) to be identified as, what we call **symplectic** coordinates. The Hamiltonian is now given by: $\pi_\mu(x)\dot{x}^\mu-L(x,\dot{x})=H(x)$. Inverting (E.32) using (E.21), Hamilton's equations are:

$$\dot{x}^\mu(t)=f^{\mu\nu}(x)\,\partial_\nu H(x)=\partial^\mu H(x)\,, \quad \text{(E.33)}$$

the solution of which defines a curve in phase space $x^\mu(t)$ starting from some initial values: $x^\mu_0:=x^\mu(0)$. The state of the system is specified by a point x on the manifold. Now let $v^\mu(x,t)$ be the flow velocity of points in phase space, defined by:

$$v^\mu(x,t)=\dot{x}^\mu(t)\,. \quad \text{(E.34)}$$

But since $\dot{x}^\mu(t)$ satisfies Hamilton's equations (E.11), the divergence of $v^\mu(x,t)$ vanishes:

$$\partial_\mu v^\mu(x,t)=\partial_\mu\dot{x}^\mu(t)=\partial_\mu\,(\,f^{\mu\nu}(x)\,\partial_\nu H(x)\,)=\Box\,H(x)\,, \quad \text{(E.35)}$$

where

$$\Box=\partial_\mu\,\partial^\mu=\partial_\mu\,f^{\mu\nu}(x)\,\partial_\nu=(\,\partial_\mu\,f^{\mu\nu}(x)\,)\,\partial_\nu\equiv 0\,. \quad \text{(E.36)}$$

where we have used Bianchi's identity [?]. This means that if we think of phase space as a fluid, the flow is such that the velocity field has no sources.

When we *can* identify canonical coordinates (q,p), a volume element $\mathrm{d}\Gamma$ in phase space is given by:

$$\mathrm{d}\Gamma = \frac{\mathrm{d}^n q\,\mathrm{d}^n p}{(2\pi\hbar)^n} = \frac{\mathrm{d}^{2n}x}{(2\pi\hbar)^n}\,. \tag{E.37}$$

For the case of general symplectic coordinates x, this volume element is described by a *volume form*, which is discussed in Section E.3.2 below. Here $\hbar$ is a factor which we introduce so as to make the phase space differential dimensionless.[3] Now let $\rho(x,t)$ be the number of states per unit volume in phase space at time t then the number of states in a region of phase space is given by:

$$N(t) = \int \mathrm{d}\Gamma\,\rho(x,t)\,, \tag{E.38}$$

where $\rho(x,t)$ is the density of states at point x in phase space at time t. Liouville's theorem states that:

Theorem 63 (Liouville's theorem)**.** *If $\rho(x,t)$ satisfies a conservation equation of the form:*

$$\frac{\partial\rho(x,t)}{\partial t} + \partial_\mu\left[\,\rho(x,t)\,v^\mu(x,t)\,\right] = 0\,, \tag{E.39}$$

and $\rho(x,t)\to 0$ as $x\to\infty$ in any direction, then the number of states $N(t)=N$ is constant in time.

Proof. We prove this by using a generalized form of Gauss' theorem.[4] Let V be a volume in phase space containing $\rho(x,t)$, and let S be the surface area of that volume. Then we find:

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = \int_V \mathrm{d}\Gamma\,\frac{\partial\rho(x,t)}{\partial t} = -\int \mathrm{d}\Gamma\,\partial_\mu\left[\,\rho(x,t)\,v^\mu(x,t)\,\right] = -\int_S \mathrm{d}S\left[\,\rho(x,t)\,v^\mu(x,t)\,\right] \to 0\,, \tag{E.40}$$

as $R\to\infty$, since $\rho(x,t)\to 0$ as $R\to\infty$. So $N(t)=N$ is a constant. □

Remark 40. Using (E.35), Eq. (E.39) becomes:

$$\frac{\partial\rho(x,t)}{\partial t} + (\,\partial_\mu\rho(x,t)\,)\,\dot{x}^\mu(t) = \frac{\partial\rho(x,t)}{\partial t} + (\,\partial_\mu\rho(x,t)\,)\,f^{\mu\nu}\,(\,\partial_\nu H(x)\,) = 0\,, \tag{E.41}$$

from which we find that $\rho(x,t)$ satisfies the equation of motion:

$$\frac{\partial\rho(x,t)}{\partial t} = -\{\,\rho(x,t), H(x)\,\}\,. \tag{E.42}$$

At $t=0$, $\rho(x,0)=\rho_0(x)$. Note that Eq. (E.42) has the *opposite sign* from Poisson's equations of motion (E.9) for the coordinates x:

$$\dot{x}^\mu(t) = \{\,x^\mu(t), H(x)\,\}\,. \tag{E.43}$$

That is, the density function moves in a time-reversed way. We will see how important this is below.

Remark 41. Since the number of states are constant under Hamiltonian flow, we can just normalize the density to unity. That is we write:

$$N = \int \mathrm{d}\Gamma\,\rho(x,t) = \int \mathrm{d}\Gamma_0\,\rho_0(x_0) = 1\,. \tag{E.44}$$

where, at $t=0$, we have set $\mathrm{d}\Gamma_0 = \mathrm{d}^{2n}x_0$ and $\rho(x_0,0)=\rho_0(x_0)$.

[3] We emphasize again that we are studying a *classical* theory here. None of our answers in this section can depend on $\hbar$. We only introduce $\hbar$ so as to make the classical phase space density dimensionless.

[4] This is called *Stokes' theorem* in geometry. We will define the integration and the terms "volume" and "surface" more precisely in Section E.3.2 below. For now, we sketch the proof using analogies from the conservation of charge in electrodynamics.

Remark 42. If $\rho_0(x)$ is of the form:

$$\rho_0(x) = \begin{cases} 1 & \text{for } x \in V, \\ 0 & \text{otherwise,} \end{cases} \tag{E.45}$$

the integration is over a finite region V in phase space. Liouville's theorem then states that this phase-space volume is preserved as a function of time t.

As one might expect, classical mechanics can best be expressed using the language of differential manifolds and differential forms. We do this in the next section.

E.2 Differential geometry

In this section we discuss concepts of differential geometry applied to classical mechanics. Unfortunately, there is a certain amount of "overhead" required to understand differential geometry, but we will find the effort well worth it. We use the definitions and notation of Schutz [**?**] where detailed discussion of the concepts of differential geometry can be found. Calculus on manifolds and differential forms are explained in several references: see for example Spivak [**?**, **?**] or Flanders [3]. Our brief account here of symplectic geometry cannot touch a more complete exposition which can be found in specialized works, such as those by Berndt [**?**] and da Silva [**?**]. We follow the classical mechanics development of Das [**?**][p. 189].

Describing physical systems in terms of geometry focuses attention on a coordinate free picture of the system in terms of topological properties of the geometry, rather than the differential equations describing the dynamics. It gives a global view of the dynamics. In addition, the system can be described in a compact notation.

Definition 43 (Manifold)**.** A manifold $\mathcal{M}$ is a set of points P where each point has a open neighborhood $\mathcal{U}$ with a one to one map h to a set x of n-tuples of real numbers: $x = (x^1, x^2, \ldots, x^n) \in \mathbb{R}^n$.

The reason for introducing manifolds rather than vector spaces are that we can study interesting global topologies using manifolds. The set of n-tuples for a neighborhood of a point $P(x)$ are called coordinates. Good coordinates are any set of linearly independent ones. We will always use good coordinates here. By constructing **charts** $(\mathcal{U}, h)$, where h is a map of $\mathcal{U} \mapsto \mathbb{R}^n$, the collection of overlapping charts, called an **atlas** can be used to describe the entire manifold $\mathcal{M}$. We will usually just define things on the neighborhood $\mathcal{U}$ of a point $P(x)$ labeled by coordinates $x \in \mathbb{R}^n$ and rely on the fact that we can extend our definitions and results to the full manifold by patching with an atlas of charts.

Definition 44 (Curves)**.** A curve $C(t)$ on a manifold is a map from a real number t to a continuous set of points $P(x)$ on the manifold. Locally the curve can be described by a set of n functions: $x^\mu = x^\mu(t)$, which are (infinitely) differentiable. That is, this is a *parametric* representation of the curve $C(t)$.

Definition 45 (Functions)**.** A function $f(P)$ on a manifold maps points $P(x)$ to real numbers. Since the point $P(x)$ can be locally described by n-coordinates x^μ, we usually just write $f(P) \equiv f(x) := f(x^1, x^2, \cdots, x^n)$.

Remark 43. Let $f(x)$ be a function defined on the manifold at point $P(x)$ described by coordinates $(x^1, x^2, \cdots, x^n)$ on a curve $C(t)$ described parametrically by parameter t. Then the derivative of $f(x)$ with respect to the parameter t on the curve $C(t)$ is given by:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \dot{x}^\mu \left(\partial_\mu f(x) \right), \qquad \text{where:} \qquad \dot{x}^\mu := \frac{\partial x^\mu}{\partial t}, \qquad \partial_\mu f(x) := \frac{\partial f(x)}{\partial x^\mu}. \tag{E.46}$$

Since the same formula holds for any function $f(x)$, the *convective* derivative operator,

$$\frac{\mathrm{d}}{\mathrm{d}t} := \dot{x}^\mu \, \partial_\mu, \tag{E.47}$$

along the path of a curve $C(t)$, defines a collection of n quantities $\dot{x}^\mu$, $\mu = 1, \cdots, n$ which is a vector, in the usual sense, pointing in a direction *tangent* to $C(t)$. A different parameterization of the same curve, say λ, where $t = t(\lambda)$, produces a new set of n quantities:

$$\frac{\partial x^\mu(\lambda)}{\partial \lambda} = \dot{x}^\mu \frac{\partial t(\lambda)}{\partial \lambda} \,, \tag{E.48}$$

which point in the same direction but are a different length, which define a different convective derivative. A different curve produces a vector pointing in some other direction. So the collection of these curves and parameters define all possible convective derivatives at point P. It is a remarkable fact that we always can find some curve $C(\lambda)$ and some parameter λ for any directional derivative operator. This leads to the following definition of vectors for manifolds:

Definition 46 (Vectors)**.** A vector[5] $\bar{v}_x$ at the point $P(x)$ is the convective derivative operator on a curve $C(t)$, parameterized by t. We write these vectors as:

$$\bar{v}_x := \frac{\mathrm{d}}{\mathrm{d}t}\Big|_x = v^\mu(x)\,\partial_\mu \,. \tag{E.49}$$

Remark 44. A vector is the convective derivative of some curve! This definition depends only on the existence of a set of curves on the manifold, and not of the local coordinate system $x \in U$ used to describe the point $P(x)$. It does not depend, for example, on transformation properties between coordinate systems. However this definition of vectors means that vectors at two points P and P' are *not* related. Additional structure on the manifold is required to compare two vectors at two different points in the manifold. We will return to this point later.

Remark 45. A particularly useful set of vectors are the convective derivatives of curves along good coordinates. That is, for a set of curves given by:

$$x^\mu(t) = t \,, \qquad \text{for all } \mu = 1, 2, \ldots, n, \tag{E.50}$$

the vectors associated with these curves, given by Eq. (E.47), are just the partial derivatives $\bar{\partial}_\mu \equiv \partial_\mu$ and provide a set of n linearly independent basis vectors.

Definition 47 (Infinitesimal vectors)**.** We define an infinitesimal vector $\bar{d}_x$ on the curve $C(t)$ at the point $P(x)$ by:

$$\bar{d}_x := \mathrm{d}t\,\frac{\mathrm{d}}{\mathrm{d}t}\Big|_x = \mathrm{d}t\,\dot{x}^\mu\,\bar{\partial}_\mu = \mathrm{d}x^\mu\,\bar{\partial}_\mu \,, \tag{E.51}$$

which is independent of the parameter t. This equation relates $\bar{d}_x$ to displacements along the curve $C(t)$ described by the coordinate displacements $\mathrm{d}x^\mu$ for the basis set $\bar{\partial}_\mu$.

Definition 48 (Tangent space)**.** It is easy to show that the convective derivatives of a collection of all curves with all parameters at a point $P(x)$ described by coordinates $x \in \mathbb{R}^n$ form a vector space for each x, called the **tangent space** $T_x\mathcal{M}$. The tangent space at x has n dimensions, the same as the manifold $\mathcal{M}$.

Exercise 82. Show that two vectors $\bar{v}$ and $\bar{w}$, defined by the derivatives:

$$\bar{v} = \frac{\mathrm{d}}{\mathrm{d}t} \,, \qquad \text{and} \qquad \bar{w} = \frac{\mathrm{d}}{\mathrm{d}s} \,, \tag{E.52}$$

on the curves $C(t)$ and $D(s)$ at point P, satisfy the requirements of a vector space, namely that $a\,\bar{v} + b\,\bar{w}$, where a and b are real-valued numbers, is some other vector at the point P, and that $\bar{v}$ and $\bar{w}$ satisfy the usual commutative and associative rules of algebra (see Section 1.1).

[5] We use an over-bar to indicate vectors.

Definition 49 (Frames)**.** Any set of n linearly independent vector fields $\bar{b}_{x\,\mu}$, $\mu = 1, 2, \ldots, n$ at point $P(x)$ described in the open neighborhood $\mathcal{U}$ with coordinates x provides a basis for the tangent vector space $T_x\mathcal{M}$ and is called a **frame**. The basis vectors for a **coordinate frame** are the partial derivatives evaluated at x:

$$\bar{b}_{x\,\mu} = \partial_\mu|_x \,, \tag{E.53}$$

and are tangent to the coordinate lines. A change of basis at point $P(x)$ in the manifold is given by:

$$\bar{b}'_{x\,\mu} = \bar{b}_{x\,\nu}\,[\,\gamma^{-1}(x)\,]^\nu{}_\mu \,, \qquad \gamma(x) \in GL_n \tag{E.54}$$

where $\gamma(x)$ is a general non-degenerate (invertible) $n \times n$ matrix. Note that a gauge group GL_n has appeared here. If the two basis vectors are coordinate bases,

$$\gamma^\mu{}_\nu(x) = \frac{\partial x'^{\,\mu}}{\partial x^\nu} \,. \tag{E.55}$$

Any vector $\bar{v}_x$ at point $P(x)$ can be expanded in either basis:

$$\bar{v}_x = v^\mu(x)\,\bar{b}_{x\,\mu} = v'^{\,\mu}(x)\,\bar{b}'_{x\,\mu} \,, \qquad v'^{\,\mu}(x) = [\,\gamma(x)\,]^\mu{}_\nu\, v^\nu(x) \,. \tag{E.56}$$

We emphasize again that vectors at different points $P(x)$ in the manifold $\mathcal{M}$ are unrelated.

Remark 46. The collection of vectors defined by a rule for selecting a vector at every point P in the manifold is called a **vector field**. A **fiber bundle** is a more general concept and consists of a base manifold with a tangent vector field, or fiber, attached to each point P in the base manifold. If the base manifold has dimension m and the vector field dimension n, then the fiber bundle is a manifold with dimension $m + n$. The fiber bundle is a special manifold, one that is decomposable by a projection from any point of a fiber onto the base manifold. So a curve on the fiber bundle constitutes a rule for assigning a vector to each point on the base manifold. Each vector on a curve in the fiber bundle is a vector field. For our case, the fiber bundle we have defined on $\mathcal{M}$ is called the **tangent bundle** $T\mathcal{M}$ and is a manifold of dimension $2n$. (See Schutz [**?**][p. 37] for further details and examples of fiber bundles.)

Definition 50 (One-forms)**.** A one-form[6] $\tilde{\phi}_x(\bar{v}_x) \in \mathbb{R}$ is a linear function on the tangent space $T_x\mathcal{M}$ at point $P(x)$ which maps a vector $\bar{v}_x$ to a real number. We require one-forms to be linear with respect to arguments:

$$\tilde{\phi}_x(\,a(x)\,\bar{v}_x + b(x)\,\bar{w}_x\,) = a(x)\,\tilde{\phi}_x(\,\bar{v}_x\,) + b(x)\,\tilde{\phi}_x(\,\bar{w}_x\,) \,, \qquad a(x), b(x) \in \mathbb{R} \,. \tag{E.57}$$

The sum and multiplication of one-forms by scalars at $P(x)$ are defined to obey:

$$(\,a(x)\,\widetilde{\phi + b(x)\,\psi}\,)_x(\bar{v}_x) = a(x)\,\tilde{\phi}_x(\bar{v}_x) + b(x)\,\tilde{\psi}_x(\bar{v}_x) \,. \tag{E.58}$$

Vectors $\bar{v}_x$ and one-forms $\tilde{\phi}_x$ belong to different spaces. One-forms are linear functions of vectors and, because of the linear requirement of one-forms, vectors can be said to be linear functions of one-forms. So we can write:

$$\tilde{\phi}_x(\bar{v}_x) \equiv \bar{v}_x(\tilde{\phi}_x) \equiv \langle\, \phi_x \,|\, v_x \,\rangle \,, \tag{E.59}$$

where in the last expression, we have used Dirac notation. That is, we can regard vectors as "kets" and one-forms as "bras," each attached to the same point $P(x)$ on $\mathcal{M}$. So the set of all one-forms at $P(x)$ also forms a vector space, called the **cotangent space**: $T^*_x\mathcal{M}$.

For a vector basis set $\{\,\bar{b}_{\mu\,x}\,\} \in T_x\mathcal{M}$, we define a corresponding one-form basis set $\{\,\tilde{b}^\mu_x\,\} \in T^*_x\mathcal{M}$ by the relation:

$$\tilde{b}^\mu_x(\,\bar{b}_{\nu\,x}\,) = \delta^\mu_\nu \,, \qquad \text{for } \mu, \nu = 1, 2, \ldots, n. \tag{E.60}$$

[6] We use a tilde to indicate a one-form.

The set of basis vectors and one-forms are said to be **dual** to each other. Any one form can be expanded $\tilde{\phi}_x$ at $P(x)$ in basis forms:

$$\tilde{\phi}_x = \phi_\mu(x)\,\tilde{b}^\mu_x\,. \tag{E.61}$$

From (E.60), we see that any one form on an arbitrary vector has the value:

$$\tilde{\phi}_x(\bar{v}_x) = \phi_\mu(x)\,v^\mu(x)\,, \tag{E.62}$$

called the **contraction** of a one-form with a vector. A change of basis in the tangent space given by Eq. (E.54) induces a corresponding basis change of the dual vectors given by:

$$\tilde{b}'^\mu_x = [\,\gamma(x)\,]^\mu{}_\nu\,\tilde{b}^\nu_x\,, \tag{E.63}$$

as can easily be checked. We will show later how to find the dual of a vector.

Definition 51 (Holonomic frames)**.** If the basis set of one-forms is exact, then we can write:

$$\tilde{b}^\mu_x = \tilde{\mathrm{d}}x^\mu\,, \tag{E.64}$$

and the frame is called **holonomic**. The dual vectors are written as $\bar{\partial}_\mu$ and obey the relation:

$$\tilde{\mathrm{d}}x^\mu(\bar{\partial}_\nu) = \delta^\mu_\nu\,. \tag{E.65}$$

Example 48. A frame given by $\tilde{b}^1 = \sin\theta\,\tilde{\mathrm{d}}\phi$, $\tilde{b}^2 = \tilde{\mathrm{d}}\theta$ is not holonomic because: $\tilde{\mathrm{d}}(\sin\theta\,\tilde{\mathrm{d}}\phi) = \cos\theta\,\tilde{\mathrm{d}}\theta\wedge\tilde{\mathrm{d}}\phi \neq 0$.

Definition 52 (Tensors)**.** We generalize one-forms and vectors to define (q,p)-tensors as fully linear functions defined on the manifold at point P which take q one-forms and p vectors as arguments and produce real-valued numbers. We write these general tensors as:

$$\mathsf{t}_x(\,\tilde{a}_x,\tilde{b}_x,\bar{c}_x,\tilde{d}_x,\bar{e}_x,\dots\,) = t^{\alpha,\beta}{}_\gamma{}^\delta{}_{\epsilon,\dots}(x)\,a_\alpha(x)\,b_\beta(x)\,c^\gamma(x)\,d_\delta(x)\,e^\epsilon(x)\,\cdots\,, \tag{E.66}$$

where we have used the linearity property, and where

$$t^{\alpha,\beta}{}_\gamma{}^\delta{}_{\epsilon,\dots}(x) = \mathsf{t}_x(\,\tilde{\mathrm{d}}x^\alpha,\tilde{\mathrm{d}}x^\beta,\bar{\partial}_\gamma,\tilde{\mathrm{d}}x^\delta,\bar{\partial}_\epsilon,\dots\,)\,. \tag{E.67}$$

The ordering of the one-forms and vectors here is important. We can only add and subtract like tensors.

Example 49. Let us construct a $(1,1)$ tensor by the **direct product**, or tensor product, of a one-form $\tilde{a}$ and a vector $\bar{b}$ which we write as: $\mathsf{t} = \tilde{a}\otimes\bar{b}$, and which has the following value when operating on a vector $\bar{c}$ and a one-form $\tilde{d}$ at point $P(x)$:

$$\mathsf{t}_x(\bar{c}_x,\tilde{d}_x) \equiv \tilde{a}_x\otimes\bar{b}_x\,(\bar{c}_x,\tilde{d}_x) \equiv \tilde{a}_x(\bar{c}_x)\,\bar{b}_x(\tilde{d}_x)\,. \tag{E.68}$$

The components of f_x in a coordinate basis are defined by:

$$t_\mu{}^\nu(x) = \mathsf{t}_x(\bar{\partial}_\mu,\tilde{\mathrm{d}}x^\nu) = a_\mu(x)\,b^\nu(x)\,, \tag{E.69}$$

so that

$$\mathsf{t}_x = \tilde{a}_x\otimes\bar{b}_x = a_\mu(x)\,b^\nu(x)\,\tilde{\mathrm{d}}x^\mu\otimes\bar{\partial}_\nu = t_\mu{}^\nu(x)\,\tilde{\mathrm{d}}x^\mu\otimes\bar{\partial}_\nu\,. \tag{E.70}$$

Example 50. A general $(0,2)$-tensor f can be written in terms of components in a coordinate basis as:

$$\mathsf{f}_x(\bar{a},\bar{b}) = f_{\mu,\nu}(x)\,\tilde{\mathrm{d}}x^\mu\otimes\tilde{\mathrm{d}}x^\nu(\bar{a},\bar{b}) = f_{\mu,\nu}(x)\,\tilde{\mathrm{d}}x^\mu(\bar{a})\,\tilde{\mathrm{d}}x^\nu(\bar{b})\,, \tag{E.71}$$

for arbitrary vectors $\bar{a}$ and $\bar{b}$. We can write this as a sum of symmetric and antisymmetric parts in the usual way. Let

$$\mathsf{f}_x(\bar{a},\bar{b}) = \mathsf{f}^{(S)}_x(\bar{a},\bar{b}) + \mathsf{f}^{(A)}_x(\bar{a},\bar{b})\,, \tag{E.72}$$

where $\mathsf{f}^{(S)}$ is even and $\mathsf{f}^{(A)}$ odd on interchange of the arguments. The symmetric part $\mathsf{f}^{(S)}$ is given by:

$$\begin{aligned}\mathsf{f}^{(S)}_x(\bar{a},\bar{b}) &= \frac{1}{2}\left[\,\mathsf{f}_x(\bar{a},\bar{b}) + \mathsf{f}_x(\bar{b},\bar{a})\,\right] \\ &= \frac{1}{2}\,f_{\mu,\nu}(x)\left[\,\tilde{\mathrm{d}}x^\mu(\bar{a})\,\tilde{\mathrm{d}}x^\nu(\bar{b}) + \tilde{\mathrm{d}}x^\mu(\bar{b})\,\tilde{\mathrm{d}}x^\nu(\bar{a})\,\right] \\ &= \frac{1}{2}\,f^{(S)}_{\mu\nu}(x)\left[\,\tilde{\mathrm{d}}x^\mu\otimes\tilde{\mathrm{d}}x^\nu + \tilde{\mathrm{d}}x^\nu\otimes\tilde{\mathrm{d}}x^\mu\,\right](\bar{a},\bar{b})\end{aligned} \tag{E.73}$$

where $f^{(S)}_{\mu\nu}(x) = [\,f_{\mu,\nu}(x) + f_{\nu,\mu}(x)\,]/2$ is even on interchange of the indices The antisymmetric part of the tensor $\mathsf{f}^{(A)}$ is given by:

$$\begin{aligned}\mathsf{f}^{(A)}_x(\bar{a},\bar{b}) &= \frac{1}{2}\left[\,\mathsf{f}_x(\bar{a},\bar{b}) - \mathsf{f}_x(\bar{b},\bar{a})\,\right] \\ &= \frac{1}{2}\,f_{\mu,\nu}(x)\left[\,\tilde{\mathrm{d}}x^\mu(\bar{a})\otimes\tilde{\mathrm{d}}x^\nu(\bar{b}) - \tilde{\mathrm{d}}x^\mu(\bar{b})\otimes\tilde{\mathrm{d}}x^\nu(\bar{a})\,\right] \\ &= \frac{1}{2}\,f^{(A)}_{\mu\nu}(x)\left[\,\tilde{\mathrm{d}}x^\mu\otimes\tilde{\mathrm{d}}x^\nu - \tilde{\mathrm{d}}x^\nu\otimes\tilde{\mathrm{d}}x^\mu\,\right](\bar{a},\bar{b})\end{aligned} \tag{E.74}$$

where $f^{(A)}_{\mu\nu}(x) = [\,f_{\mu,\nu}(x) - f_{\nu,\mu}(x)\,]/2$ is odd on interchange of the indices.

Definition 53 (Metric tensors)**.** A metric tensor g_x at point $P(x)$ in the manifold is a *non-singular* and *symmetric* $(0,2)$-tensor, not a form.[7] The metric is usually also required to be positive definite. For coordinates x at point $P(x)$ in the manifold, it is defined in a general frame by:

$$\mathsf{g}_x = g_{\mu\nu}(x)\,\tilde{b}^\mu_x\otimes\tilde{b}^\nu_x\,,\qquad \text{where}\quad g_{\mu\nu}(x)\equiv\mathsf{g}_x(\bar{b}_{x\,\mu},\bar{b}_{x\,\nu})\,, \tag{E.75}$$

which takes two vectors as arguments. The symmetry requirement means that $g_{\mu\nu}(x) = g_{\nu\mu}(x)$. The non-singular requirement means that $g_{\mu\nu}(x)$ is invertible. We write the inverse as $g^{\mu\nu}(x)$ and obtain:

$$g^{\mu\nu}(x)\,g_{\nu\lambda}(x) = \delta^\mu_\lambda\,, \tag{E.76}$$

for all points $P(x)$.

Remark 47. The length $\|\mathrm{d}\ell_x\|$ in a coordinate frame of the infinitesimal vector $\bar{d}_x$ defined in Eq. (E.51) is given by:

$$\|\mathrm{d}\ell_x\|^2 = \mathsf{g}_x(\bar{d}_x,\bar{d}_x) = g_{\mu\nu}(x)\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu\,. \tag{E.77}$$

If there is a metric defined on the manifold, we can use it to relate vectors to corresponding forms. That is, if $\bar{v}_x\in T_x\mathcal{M}$, then it's dual one-form $\tilde{v}_x\in T^*_x\mathcal{M}$ is given by:

$$\tilde{v}_x = \mathsf{g}_x(\bar{v}_x) = g_{\mu\nu}(x)\,v^\mu(x)\,\tilde{b}^\nu \equiv v_\nu(x)\,\tilde{b}^\nu\,, \tag{E.78}$$

so that the components of the one-form $\tilde{v}_x$ are given by:

$$v_\mu(x) = g_{\mu\nu}(x)\,v^\nu(x)\,,\qquad v^\mu(x) = g^{\mu\nu}(x)\,v_\nu(x)\,, \tag{E.79}$$

where we have used the symmetry property of the metric and the inverse relation (E.76).

Remark 48. A transformation to a new coordinate frame, given by Eq. (E.63):

$$\tilde{b}'^\mu_x = [\,\gamma(x)\,]^\mu{}_\nu\,\tilde{b}^\nu_x\,,\qquad \gamma(x)\in GL_n\,. \tag{E.80}$$

Since the metric tensor g_x is *independent* of the frame, the change to a new coordinate frame means that the matrix of the metric tensor transforms according to:

$$g'_{\mu\nu}(x') = g_{\mu'\nu'}(x)\,[\,\gamma^{-1}(x)\,]^{\mu'}{}_\mu\,[\,\gamma^{-1}(x)\,]^{\nu'}{}_\nu \tag{E.81}$$

Coordinate frame transformations $\gamma(x)$ which preserve the metric (such as matrices which belong to the Lorentz group in flat space-time), are called **isometries**.

[7] Recall that a form is an *anti*-symmetric $(0,2)$-tensor.

Remark 49. An **orthonormal** frame is one in which:

$$g_{\mu\nu}(x) = \delta_{\mu,\nu}\,, \tag{E.82}$$

and is independent of x. If a frame is both holonomic and orthonormal the coordinates are called **Cartesian**. The metric is then called **Euclidean**. A four-dimensiona **Minkowsky metric**, used in relativity, is given by:

$$g_{\mu\nu}(x) \equiv \eta_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)\,. \tag{E.83}$$

We usually reserve the symbol $\eta_{\mu\nu}$ for the metric, and write:[8] The Minkowsky metric is not positive definite and for this reason is sometimes called a **pseudo-metric**. There are no differences between upper and lower indices for manifolds with Euclidean metrics: $v_\mu = g_{\mu\nu}\,v^\nu = v^\mu$. The space-parts of a vector change sign for a Minkowsky metric. Matrices which belong to the Lorentz group are isometric transformation of the Minkowsky metric.

Definition 54 (p-forms)**.** We note that one-forms, which take a vector into a real number, are the same as our definition of a $(0,1)$-tensor. We will find it useful to define p-forms to be fully *antisymmetric* $(0,p)$-tensors which take p-vectors into a real number. We write p-forms with a tilde. The antisymmetry of a p-form means that for all i and j:

$$\tilde{\phi}_x(\bar{v}_1,\ldots,\bar{v}_i,\ldots,\bar{v}_j,\ldots,\bar{v}_p) = -\tilde{\phi}_x(\bar{v}_1,\ldots,\bar{v}_j,\ldots,\bar{v}_i,\ldots,\bar{v}_p)\,. \tag{E.84}$$

The space of p-forms is written as: $\Lambda^p T_x\mathcal{M}$. So in this notation, the **cotangent** space of one-forms is the same as $T_x^*\mathcal{M} \equiv \Lambda^1 T_x\mathcal{M}$. By convention, zero-forms are functions: $\Lambda^0 T_x\mathcal{M} = \mathbb{R}$. Zero-forms acting on vectors are defined to be zero. There are no p-forms for $p > n$, the dimension of the manifold.

We write $\Lambda\mathcal{M}$ as the direct sum of the collection of p-forms at $P(x)$:

$$\Lambda\mathcal{M} = \bigoplus_{p=0}^{n} \Lambda^p\mathcal{M}\,. \tag{E.85}$$

Because of the antisymmetry of p-forms, the dimension of $\Lambda^p\mathcal{M}$ and $\Lambda\mathcal{M}$ are given by:

$$\dim\{\,\Lambda^p\mathcal{M}\,\} = \binom{n}{p} = \frac{n!}{p!\,(n-p)!}\,, \qquad \dim\{\,\Lambda\mathcal{M}\,\} = 2^{2n}\,, \tag{E.86}$$

so that the dimension of ΛM is even, for $n \geq 1$.

Definition 55 (Wedge product)**.** Let $\tilde{\phi}_x$ be a p-form and $\tilde{\psi}_x$ be a q-form. Then the wedge product is a $(p+q)$-form given by:

$$(\tilde{\phi}_x \wedge \tilde{\psi}_x)(\,\bar{v}_1,\ldots,\bar{v}_{p+q}\,) := \frac{1}{p!q!}\sum_\pi (-)^\pi \tilde{\phi}_x(\,\bar{v}_{\pi(1)},\ldots,\bar{v}_{\pi(p)}\,)\,\tilde{\psi}_x(\,\bar{v}_{\pi(p+1)},\ldots,\bar{v}_{\pi(p+q)}\,)\,, \tag{E.87}$$

where π runs over all permutations of $p+q$ objects. For example, the wedge product of two one-forms $\tilde{a}$ and $\tilde{b}$ at point $P(x)$ is given by:

$$\begin{aligned}\tilde{a}\wedge\tilde{b}\,(\,\bar{v}_1,\bar{v}_2\,) &= \tilde{a}(\bar{v}_1)\,\tilde{b}(\bar{v}_2) - \tilde{a}(\bar{v}_2)\,\tilde{b}(\bar{v}_1)\\ &= \tilde{a}(\bar{v}_1)\,\tilde{b}(\bar{v}_2) - \tilde{b}(\bar{v}_1)\,\tilde{a}(\bar{v}_2)\\ &= \left\{\,\tilde{a}\otimes\tilde{b} - \tilde{b}\otimes\tilde{a}\,\right\}(\,\bar{v}_1,\bar{v}_2\,)\,.\end{aligned} \tag{E.88}$$

So we can just write:

$$\tilde{a}\wedge\tilde{b} = \tilde{a}\otimes\tilde{b} - \tilde{b}\otimes\tilde{c}\,. \tag{E.89}$$

[8] This is the "particle physicists" metric. The one used most often in general relativity is with the signature: $(-1,1,1,1)$.

The wedge product of three one-forms is the fully antisymmetric combination:

$$\begin{aligned}\tilde{a}\wedge\tilde{b}\wedge\tilde{c} &= (\tilde{a}\wedge\tilde{b})\wedge\tilde{c} = \tilde{a}\wedge(\tilde{b}\wedge\tilde{c}) \\ &= \tilde{a}\otimes\tilde{b}\otimes\tilde{c}+\tilde{b}\otimes\tilde{c}\otimes\tilde{a}+\tilde{c}\otimes\tilde{a}\otimes\tilde{b}-\tilde{b}\otimes\tilde{a}\otimes\tilde{c}-\tilde{c}\otimes\tilde{b}\otimes\tilde{a}-\tilde{a}\otimes\tilde{c}\otimes\tilde{b}\end{aligned} \tag{E.90}$$

Now let $\tilde{b}^{\mu}_{x}$ be a basis form at $P(x)$. Then a p-form can be written as:

$$\tilde{\phi}_x = \frac{1}{p!}\sum_{\mu_1,\ldots,\mu_p}\phi_{\mu_1,\ldots,\mu_p}(x)\,\tilde{b}^{\mu_1}_{x}\wedge\tilde{b}^{\mu_2}_{x}\wedge\cdots\wedge\tilde{b}^{\mu_p}_{x}\,, \tag{E.91}$$

where $\phi_{\mu_1,\ldots,\mu_p}(x)$ is fully antisymmetric in all values of indices.

If $\tilde{\phi}_n$ are a p-forms and $\tilde{\psi}_n$ are q-forms, then the wedge product obeys the following rules:

1. bilinear:

$$\begin{aligned}(a\,\tilde{\phi}_1+b\,\tilde{\phi}_2)\wedge\tilde{\psi} &= a\,\tilde{\phi}_1\wedge\tilde{\psi}+b\,\tilde{\phi}_2\wedge\tilde{\psi}\,, \qquad (a,b)\in\mathbb{R}\,, \\ \tilde{\phi}\wedge(c\,\tilde{\psi}_1+d\,\tilde{\psi}_2) &= c\,\tilde{\phi}\wedge\tilde{\psi}_1+d\,\tilde{\phi}\wedge\tilde{\psi}_2\,, \qquad (c,d)\in\mathbb{R}\,.\end{aligned} \tag{E.92}$$

2. associative:

$$(\tilde{\phi}\wedge\tilde{\psi})\wedge\tilde{\chi} = \tilde{\phi}\wedge(\tilde{\psi}\wedge\tilde{\chi})\,. \tag{E.93}$$

3. graded commutative:

$$\tilde{\phi}\wedge\tilde{\psi} = (-)^{pq}\,\tilde{\psi}\wedge\tilde{\phi}\,. \tag{E.94}$$

In particular, $\tilde{\phi}\wedge\tilde{\phi}=0$ if p is odd.

Definition 56 (Contraction)**.** The **contraction** of a vector $\bar{v}$ at point $P(x)$ with a p-form $\tilde{\phi}$ produces $(p-1)$-form $\tilde{\psi}$. The contraction is usually defined by putting the vector into the first slot of $\tilde{\phi}$:[9]

$$\tilde{\psi} := \tilde{\phi}(\bar{v}) = \tilde{\phi}(\underbrace{\bar{v},\cdot,\cdot,\cdots}_{p\text{ slots}}) = \phi_{\mu_1,\mu_2,\ldots}(x)\,v^{\mu_1}(x)\,\tilde{\mathrm{d}}x^{\mu_2}\wedge\cdots\equiv\psi_{\mu 2,\sigma\ldots}(x)\,\tilde{\mathrm{d}}x^{\mu_2}\wedge\cdots\,. \tag{E.95}$$

E.3 The calculus of forms

E.3.1 Derivatives of forms

In this section we define interior and exterior derivatives of forms.

Definition 57 (Interior derivative)**.** The interior derivative $\mathrm{i}_{\bar{v}}$ at point $P(x)$ of a p-form $\tilde{\omega}$ with respect to any vector $\bar{v}$ is a contraction which reduces $\tilde{\omega}$ to a $(p-1)$-form. We write:[10]

$$\mathrm{i}_{\bar{v}}\,\tilde{\omega} := \tilde{\omega}(\bar{v})\,. \tag{E.96}$$

Remark 50. Interior derivatives obey the following rules:

1. Linearity in $\bar{v}$:

$$\mathrm{i}_{a\bar{v}+b\bar{w}} = a\,\mathrm{i}_{\bar{v}}+b\,\mathrm{i}_{\bar{w}}\,, \qquad (a,b)\in\mathbb{R}\,. \tag{E.97}$$

2. The Leibniz rule: Let $\tilde{\omega}$ be a p-form and $\tilde{\sigma}$ a q-form. Then:

$$\mathrm{i}_{\bar{v}}\,(\tilde{\omega}\wedge\tilde{\sigma}) = (\,\mathrm{i}_{\bar{v}}\tilde{\omega}\,)\wedge\tilde{\sigma}+(-)^{p}\,\tilde{\omega}\wedge(\,\mathrm{i}_{\bar{v}}\tilde{\sigma}\,)\,. \tag{E.98}$$

[9] However, it any slot is OK.

[10] Loomis and Sternberg [?][p. 456] use the symbol $\lrcorner$ to denote the interior derivative: $\bar{v}\lrcorner\tilde{\omega}\equiv\mathrm{i}_{\bar{v}}\,\tilde{\omega}$.

3. $\mathrm{i}_{\bar{v}}\,\mathrm{i}_{\bar{w}} + \mathrm{i}_{\bar{w}}\,\mathrm{i}_{\bar{v}} = 0$.

By definition, the interior derivative of a zero form (a function) is zero: $\mathrm{i}_{\bar{v}}(f(x)) = 0$.

The word "derivative" refers here to a purely algebraic property of the interior derivative, which satisfy the rules given above. This definition is an example of a more general property called a "derivation" (for more details, see Göckeler and Schücke [4][p. 86]).

Definition 58 (Gradient-form)**.** The gradient one-form of a function $f(x)$ defined on $\mathcal{M}$, is defined to be a one-form $\tilde{\mathrm{d}}f(x)$. Evaluated on an arbitrary vector $\bar{v}_x = \mathrm{d}/\mathrm{d}t|_x$, it is defined by the rule:

$$\tilde{\mathrm{d}}f(x)(\bar{v}_x) := \frac{\mathrm{d}f}{\mathrm{d}t}\Big|_x . \tag{E.99}$$

In particular, setting $f(t) = x^\mu(t)$, we find:

$$\tilde{\mathrm{d}}x^\mu(\bar{v}_x) = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}\Big|_x = \dot{x}^\mu , \tag{E.100}$$

so that (E.99) becomes:

$$\tilde{\mathrm{d}}f(x)(\bar{v}_x) = (\partial_\mu f(x))\,\dot{x}^\mu = (\partial_\mu f(x))\,\tilde{\mathrm{d}}x^\mu(\bar{v}_x) , \tag{E.101}$$

for any vector $\bar{v}_x$. So we can just write the one-form $\tilde{\mathrm{d}}f(x)$ as:

$$\tilde{\mathrm{d}}f(x) := (\partial_\mu f(x))\,\tilde{\mathrm{d}}x^\mu . \tag{E.102}$$

Remark 51. If $\bar{v}_x$ is one of the basis vectors $\bar{\partial}_\nu$ along a coordinate line, then it is easy to see that our definition of the gradient form means that the basis set of one-forms $\{\tilde{\mathrm{d}}x^\mu\}$ are dual to the basis set of coordinates $\{\bar{\partial}_\mu\}$ at point $P(x)$:

$$\tilde{\mathrm{d}}x^\mu(\,\bar{\partial}_\nu\,) = \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu . \tag{E.103}$$

Remark 52. From the definition of $\bar{d}_x$ in Eq. (E.51), we find that:

$$\tilde{\mathrm{d}}x^\mu(\,\bar{d}_x\,) = \mathrm{d}x^\nu\,\tilde{\mathrm{d}}x^\mu(\,\bar{\partial}_\nu\,) = \mathrm{d}x^\nu\,\delta^\mu_\nu = \mathrm{d}x^\mu , \tag{E.104}$$

The gradient of a function $f(x)$ defined on $\mathcal{M}$ is a one-form, not a vector. When operating on the vector $\bar{d}_x$ at point $P(x)$, we find:

$$\tilde{\mathrm{d}}f(x)(\,\bar{d}_x\,) = (\,\partial_\mu f(x)\,)\,\mathrm{d}x^\mu , \tag{E.105}$$

which is the usual definition of the gradient operator when acting on functions.

Definition 59 (Exterior derivative)**.** In Eq. (E.102), we defined the exterior derivative of a fuction $f(x)$ by:

$$\tilde{\mathrm{d}}f(x) := (\partial_\mu f(x))\,\tilde{\mathrm{d}}x^\mu . \tag{E.106}$$

That is, the exterior derivative of a zero-form is a one-form. We wish to extend this definition to higher forms so that the exterior derivative of a p-form will raise the form to a $(p+1)$-form. We do this by the following rules:

1. Linearity in $\tilde{\omega}$: if $\tilde{\omega}$ and $\tilde{\sigma}$ are p-forms:

$$\tilde{\mathrm{d}}\,(\,a\,\tilde{\omega} + b\,\tilde{\sigma}\,) = a\,\tilde{\mathrm{d}}\,\tilde{\omega} + b\,\tilde{\mathrm{d}}\,\tilde{\sigma} , \qquad (a,b) \in \mathbb{R} . \tag{E.107}$$

2. The Leibniz rule: let $\tilde{\omega}$ be a p-form and $\tilde{\sigma}$ a q-form. Then:

$$\tilde{\mathrm{d}}\,(\,\tilde{\omega} \wedge \tilde{\sigma}\,) = (\,\tilde{\mathrm{d}}\,\tilde{\omega}\,) \wedge \tilde{\sigma} + (-)^p\,\tilde{\omega} \wedge \tilde{\mathrm{d}}\tilde{\sigma} . \tag{E.108}$$

3. $\tilde{\mathrm{d}}\,(\,\tilde{\mathrm{d}}\,\tilde{\omega}\,) = 0$, where $\tilde{\omega}$ is a p-form.

The exterior derivative of a general p-form is given in Exercise 85 below.

Example 51. Let us find the exterior derivative of a one-form using these rules. Let $\tilde{a} = a_\nu(x)\,\tilde{\mathrm{d}}x^\nu$ be a one-form. Then

$$\begin{aligned}\tilde{\mathrm{d}}\,\tilde{a} &= (\tilde{\mathrm{d}}\,a_\nu(x)) \wedge \tilde{\mathrm{d}}x^\nu + a_\nu(x)\,\tilde{\mathrm{d}}\,(\,\tilde{\mathrm{d}}x^\nu\,) = (\partial_\mu a_\nu(x))\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu \\ &= \frac{1}{2}\,\big[\,\partial_\mu a_\nu(x) - \partial_\nu a_\mu(x)\,\big]\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,.\end{aligned} \tag{E.109}$$

Exercise 83. By using components, show that $\tilde{\mathrm{d}}\tilde{\mathrm{d}}\,f(x) = 0$, where $f(x)$ is a zero form. If $f(x)$ and $g(x)$ are zero-forms, show that $\tilde{\mathrm{d}}\,(\,f(x)\,(\,\tilde{\mathrm{d}}\,g(x)\,)\,) = (\,\tilde{\mathrm{d}}f(x)\,) \wedge (\,\tilde{\mathrm{d}}g(x)\,)$.

Exercise 84. Compute $\tilde{\mathrm{d}}(\tilde{\omega}(\bar{V}))$ and $(\tilde{\mathrm{d}}\,\tilde{\omega})\,(\bar{V})$ and compare the two results.

Exercise 85. If $\tilde{f}$ is a p-form given by:

$$\tilde{f} = \frac{1}{p!}\,f_{\alpha,\beta,\ldots,\gamma}(x)\,\tilde{\mathrm{d}}x^\alpha \wedge \tilde{\mathrm{d}}x^\beta \wedge \cdots \wedge \tilde{\mathrm{d}}x^\gamma\,, \tag{E.110}$$

show that $\tilde{g}(x) = \tilde{\mathrm{d}}\,\tilde{f}(x)$ is a $(p+1)$-form given by:

$$\tilde{g} = \tilde{\mathrm{d}}\,\tilde{f} = \frac{1}{(p+1)!}\,g_{\alpha,\beta,\ldots,\gamma}(x)\,\tilde{\mathrm{d}}x^\alpha \wedge \tilde{\mathrm{d}}x^\beta \wedge \cdots \wedge \tilde{\mathrm{d}}x^\gamma\,, \tag{E.111}$$

where $g_{\alpha,\beta,\ldots,\delta}(x)$ is the antisymmetric combination:

$$g_{\alpha,\beta,\ldots,\delta}(x) = \partial_{[\alpha} f_{\beta,\gamma,\ldots,\delta]}(x)\,. \tag{E.112}$$

We also will need to know the meaning of closed and exact forms, and understand Poincaré's lemma. We define them here:

Definition 60 (Closed form)**.** If $\tilde{\omega}$ is a p-form and $\tilde{\mathrm{d}}\,\tilde{\omega} = 0$, then $\tilde{\omega}$ is said to be closed.

Definition 61 (Exact form)**.** If $\tilde{\omega}$ is a p-form which can be written as $\tilde{\omega} = \tilde{\mathrm{d}}\,\tilde{\sigma}$, where $\tilde{\sigma}$ is a $(p-1)$-form, then $\tilde{\omega}$ is called exact.

It is clear from property 3 above that an exact form is closed. The reverse is generally not true, as a simple example can show; however, if the domain of definition of the p-form is restricted to certain regions, one can show that a closed form is also exact, except for the addition of a "gauge" term. The conditions for which this is true is called **Poincaré's Lemma**.

Theorem 64 (Poincaré's Lemma)**.** *If a region $\mathcal{U}$ of the manifold is "star-shaped", and a p-form $\tilde{\omega}$ is closed in this region, it is exact.*

Proof. For a proof of the theorem, see Göckeler and Schücke [4][p. 21] or Schutz [**?**][p. 138]. □

In order to compare forms and vectors at two different points in M, we use the Lie derivative.

Definition 62 (Lie derivative)**.** The Lie derivative $\pounds_{\bar{v}}(\tilde{\omega})$ of a p-form $\tilde{\omega}$ with respect to a vector field $\bar{v}$ is defined by:

$$\pounds_{\bar{v}}(\tilde{\omega}) := (\,\tilde{\mathrm{d}}\,\mathrm{i}_{\bar{v}} + \mathrm{i}_{\bar{v}}\,\tilde{\mathrm{d}}\,)\,(\,\tilde{\omega}\,) = \tilde{\mathrm{d}}(\,\tilde{\omega}(\bar{v})\,) + (\,\tilde{\mathrm{d}}\,\tilde{\omega}\,)(\bar{v})\,. \tag{E.113}$$

By definition, the contraction of a zero-form $f_x = f(x)$ on a vector $\bar{v}$ vanishes, $f_x(\bar{v}) = 0$, so that the Lie derivative of a function (zero-form) with respect to a vector is given by:

$$\pounds_{\bar{v}}(f) = (\,\tilde{\mathrm{d}}\,f(x)\,)(\bar{v}) = (\,\partial_\mu f(x)\,)\,v^\mu\,. \tag{E.114}$$

The volume form for a manifold is a special form which we will need for defining integration of forms.

Definition 63 (Volume form)**.** The volume form for an n-dimensional manifold is an n-form defined by:

$$\tilde{\omega} = \tilde{\mathrm{d}}x^1 \wedge \tilde{\mathrm{d}}x^2 \wedge \cdots \wedge \tilde{\mathrm{d}}x^n \,. \tag{E.115}$$

Then $\tilde{\mathrm{d}}\,\tilde{\omega} = 0$, since the space is only n-dimensional.

Remark 53. Now let us suppose we change to new coordinates $X^\mu = X^\mu(x)$. Then since

$$\tilde{\mathrm{d}}x^\mu = \frac{\partial x^\mu}{\partial X^\nu}\,\tilde{\mathrm{d}}X^\mu \,, \tag{E.116}$$

we find:

$$\begin{aligned} \tilde{\omega} &= \frac{\partial x^1}{\partial X^{\mu_1}}\,\frac{\partial x^2}{\partial X^{\mu_2}} \cdots \frac{\partial x^n}{\partial X^{\mu_n}}\,\tilde{\mathrm{d}}X^{\mu_1} \wedge \tilde{\mathrm{d}}X^{\mu_2} \wedge \cdots \wedge \tilde{\mathrm{d}}X^{\mu_n} \\ &= \det\Big[\,\frac{\partial x^\mu}{\partial X^\nu}\,\Big]\,\tilde{\mathrm{d}}X^1 \wedge \tilde{\mathrm{d}}X^2 \wedge \cdots \wedge \tilde{\mathrm{d}}X^n = \det\Big[\,\frac{\partial x^\mu}{\partial X^\nu}\,\Big]\,\tilde{\Omega}\,, \\ \text{where:}\quad \tilde{\Omega} &= \tilde{\mathrm{d}}X^1 \wedge \tilde{\mathrm{d}}X^2 \wedge \cdots \wedge \tilde{\mathrm{d}}X^n \,. \end{aligned} \tag{E.117}$$

So the volume form transforms by multiplication of the form with the Jacobian of the transformation. Thus it is suitable as a volume element for integration, as we show below.

Definition 64 (Orientation of n-forms)**.** Any n ordered basis forms defines an orientation by means of a volume form. Any other volume form obtained by a change of coordinates is said to have the same orientation if the determinant of the Jacobian relating these forms is positive definite. Not every surface is orientable. For example, a Möbius strip is *not* orientable.

It is useful also to have a definition for the divergence of a vector.

Definition 65 (Divergence)**.** The divergence $\tilde{\nabla}_{\tilde{\omega}}(\bar{v})$ of a vector field $\bar{v}$ with respect to the volume form $\tilde{\omega}$ is defined by the equation:

$$\tilde{\nabla}_{\tilde{\omega}}(\bar{v})\,\tilde{\omega} \equiv \tilde{\mathrm{d}}(\,\tilde{\omega}(\bar{v})\,) \,. \tag{E.118}$$

This states that if $\tilde{\omega}$ is an n-form, then $\tilde{\omega}(\bar{v})$ is an $(n-1)$-form, the exterior derivative of which is an n-form again. This form must be proportional to $\tilde{\omega}$ again, the factor of proportionality is the divergence.

Example 52. We illustrate our definition for the case of ordinary three-dimensional Euclidean manifold, where the volume form $\tilde{\omega}$ and an arbitrary vector $\bar{v}$ is given by:

$$\begin{aligned} \tilde{\omega} &= \tilde{\mathrm{d}}x \wedge \tilde{\mathrm{d}}y \wedge \tilde{\mathrm{d}}z \,, \\ \bar{v} &= v_x\,\bar{\partial}_x + v_y\,\bar{\partial}_y + v_z\,\bar{\partial}_z \,, \end{aligned} \tag{E.119}$$

where v_x, v_y, and v_z depend on x, y, and z. So we find:

$$\tilde{\omega}(\,\bar{v}\,) = v_x\,\tilde{\mathrm{d}}y \wedge \tilde{\mathrm{d}}z - v_y\,\tilde{\mathrm{d}}x \wedge \tilde{\mathrm{d}}z + v_z\,\tilde{\mathrm{d}}x \wedge \tilde{\mathrm{d}}y \,, \tag{E.120}$$

and

$$\begin{aligned} \tilde{\mathrm{d}}(\,\tilde{\omega}(\bar{v})\,) &= (\,\partial_x v_x\,)\,\tilde{\mathrm{d}}x \wedge \tilde{\mathrm{d}}y \wedge \tilde{\mathrm{d}}z - (\,\partial_y v_y\,)\,\tilde{\mathrm{d}}y \wedge \tilde{\mathrm{d}}x \wedge \tilde{\mathrm{d}}z + (\,\partial_z v_z\,)\,\tilde{\mathrm{d}}z \wedge \tilde{\mathrm{d}}x \wedge \tilde{\mathrm{d}}y \\ &= (\,\partial_x v_x + \partial_y v_y + \partial_z v_z\,)\,\tilde{\mathrm{d}}x \wedge \tilde{\mathrm{d}}y \wedge \tilde{\mathrm{d}}z \equiv (\,\boldsymbol{\nabla}\cdot\mathbf{v}\,)\,\tilde{\omega} \,, \end{aligned} \tag{E.121}$$

so that $\tilde{\nabla}_{\tilde{\omega}}(\bar{v}) = \boldsymbol{\nabla}\cdot\mathbf{v}$, in agreement with the usual definition of the divergence of a vector in Cartesian coordinates.

E.3.2 Integration of forms

Integration of forms over regions of a manifold can be defined without a definition of a metric, or length. Nice discussions of integration can be found in Schutz [**?**][p. 144] and in Göckeler and Schücke [4][p. 22].

The point here is that the volume for a region of a manifold can be defined without the use of a metric. In fact, a form is the natural way to define volume. Let us first define the integral of an n-form (the volume form) over an oriented region U on TM. We do this in the following definition:

Definition 66 (Integration of n-forms)**.** The integral of an n-form over a volume U on TM is defined to be:

$$\int_U \tilde{\omega} := \iint \cdots \int_U \mathrm{d}x_1 \, \mathrm{d}x_2 \cdots \mathrm{d}x_n \,, \tag{E.122}$$

where, on the left-hand side we define the integral of an n- form, and on the right-hand side we have the ordinary integral of calculus.

Remark 54. This definition implies that the integral of a form is *independent* of the coordinate system used. We can easily prove this by noting from Eq. (E.117), that two forms with the same orientation are related by:

Theorem 65 (Stokes' theorem)**.** *Stokes' theorem states that the integral of the exterior derivative of a p-form $\tilde{\omega}$ over a region U on a manifold is given by the integral of $\tilde{\omega}$ evaluated on the boundary ∂U of the region. That is:*

$$\int_U \tilde{\mathrm{d}}\, \tilde{\omega} = \int_{\partial U} \tilde{\omega} \,, \tag{E.123}$$

where on the right-hand side, $\tilde{\omega}$ is restricted to the boundary ∂U.[11]

Proof. Schutz [**?**][p. 144] gives a geometric proof for $p = n - 1$, that is $\tilde{\mathrm{d}}\tilde{\omega}$ is an n-form, using Lie-dragging. The theorem can be proved for $p < n - 1$ by defining oriented sub-manifolds (See Göckeler and Schücker [4][p. 26]). □

E.4 Non-relativistic space-time

Classical mechanics is described by a fiber bundle structure with time as a one-dimensional base manifold with a symplectic manifold attached at each point in time. We illustrate this in the figure.

E.4.1 Symplectic manifolds

In this section we study properties of symplectic manifolds.

Definition 67 (Symplectic manifold)**.** A **symplectic** manifold $\mathcal{M}_S$ is one that has associated with it a *non-degenerate and closed* two-form $\tilde{f}$.

Since the two-form $\tilde{f}$ is an antisymmetric $(0, 2)$-tensor, the dimension of the manifold must be even: $n \mapsto 2n$. In a coordinate basis, we write the two-form $\tilde{f}_x$ as:

$$\tilde{f}_x = \frac{1}{2} f_{\mu\nu}(x) \, \tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu \,, \tag{E.124}$$

with $f_{\mu\nu}(x) = -f_{\nu\mu}(x)$. The statement that $\tilde{f}_x$ is non-degenerate means that $\det\{f_{\mu\nu}(x)\} \neq 0$, so that the inverse of the matrix $f_{\mu\nu}(x)$ exists. We define the inverse matrix with upper indices:

$$f^{\mu\lambda}(x) \, f_{\lambda\nu}(x) = \delta^\mu_\nu = f_{\nu\lambda}(x) \, f^{\lambda\mu}(x) \,. \tag{E.125}$$

[11] The boundary of the region U must divide the space into an "inside" and an "outside." That is, not around a wormhole.

Now let $\bar{v} = v^\mu(x)\,\bar{\partial}_\mu$ be any vector. Then we define it's dual by putting the vector $\bar{v}$ in the *second* slot[12] of the symplectic two-form $\tilde{f}_x$. That is:

$$\tilde{v}_x(\cdot) := \tilde{f}_x(\cdot, \bar{v}) = f_{\mu\nu}(x)\,\tilde{\mathrm{d}}x^\mu(\cdot)\,v^\nu(x)\,, \tag{E.126}$$

so since $\tilde{v}_x(\cdot) = v_\mu(x)\,\tilde{\mathrm{d}}x^\mu(\cdot)$, we find that:

$$v_\mu(x) = f_{\mu\nu}(x)\,v^\nu(x)\,, \qquad v^\mu(x) = f^{\mu\nu}(x)\,v_\nu(x)\,. \tag{E.127}$$

where we have used Eq. (E.125). Note that unlike the *symmetric* metric of relativity the ordering of indices here are important. The set of base one-forms $\tilde{\mathrm{d}}x_\mu$ with *lower* indices, are then defined by:

$$\tilde{\mathrm{d}}x_\mu = \tilde{\mathrm{d}}x^\nu\,f_{\nu\mu}(x)\,, \qquad \tilde{\mathrm{d}}x^\mu = \tilde{\mathrm{d}}x_\nu\,f^{\nu\mu}(x)\,, \tag{E.128}$$

so that

$$\tilde{v}_x = v_\mu(x)\,\tilde{\mathrm{d}}x^\mu = v^\mu(x)\,\tilde{\mathrm{d}}x_\mu\,. \tag{E.129}$$

where we have used Eq. (E.125). Note that $\tilde{\mathrm{d}}x_\mu(\cdot) = \tilde{f}_x(\cdot, \bar{\partial}_\mu)$. The set of base vectors $\bar{\partial}^\mu$ with *upper* indices are then given by:

$$\bar{\partial}^\mu = f^{\mu\nu}(x)\,\bar{\partial}_\nu\,, \qquad \bar{\partial}_\mu = f_{\mu\nu}(x)\,\bar{\partial}^\nu\,, \tag{E.130}$$

so that $\bar{\partial}^\mu$ and $\tilde{\mathrm{d}}x_\mu$ are duals and obey the orthogonal relation:

$$\tilde{\mathrm{d}}x_\mu(\bar{\partial}^\nu) = \delta_\mu{}^\nu\,. \tag{E.131}$$

These definitions make it easy to write vectors and one-forms as:

$$\begin{aligned} \bar{v}_x &= v^\mu(x)\,\bar{\partial}_\mu \quad = v_\mu(x)\,\bar{\partial}^\mu\,, \\ \tilde{v}_x &= v_\mu(x)\,\tilde{\mathrm{d}}x^\mu = v^\mu(x)\,\tilde{\mathrm{d}}x_\mu\,. \end{aligned} \tag{E.132}$$

Definition 68 (Hamiltonian vector field)**.** If the Lie derivative of the symplectic two-form $\tilde{f}_x$ with respect to a vector field $\bar{v}_x$ vanishes, we call $\bar{v}_x$ a Hamiltonian vector field. Since $\tilde{f}_x$ is closed, this means that a Hamiltonian vector field $\bar{v}_x$ satisfies:

$$\pounds_{\bar{v}_x}(\tilde{f}_x) = \tilde{\mathrm{d}}(\,\tilde{f}_x(\cdot, \bar{v}_x)\,) = 0\,. \tag{E.133}$$

Example 53. Let $\tilde{v}_x = \tilde{\mathrm{d}}H(x)$, where $H(x)$ is a function on $\mathcal{M}$ at point $P(x)$. Then using our definitions of upper and lower basis one-forms and vectors, we have:

$$\tilde{v}_x = (\partial_\mu H(x))\,\tilde{\mathrm{d}}x^\mu\,, \qquad \bar{v}_x = (\partial_\mu H(x))\,\bar{\partial}^\mu\,. \tag{E.134}$$

So

$$\pounds_{\bar{v}_x}(\tilde{f}_x) = \tilde{\mathrm{d}}(\,\tilde{f}_x(\cdot, \bar{v}_x)\,) = \tilde{\mathrm{d}}(\,\tilde{v}_x\,) = \tilde{\mathrm{d}}\,\tilde{\mathrm{d}}H(x) = 0\,, \tag{E.135}$$

and $\tilde{v}_x$ is a Hamiltonian vector field. Conversely, if $\bar{v}_x$ is a Hamiltonian vector field, then

$$\pounds_{\bar{v}_x}(\tilde{f}_x) = \tilde{\mathrm{d}}(\,\tilde{f}_x(\cdot, \bar{v}_x)\,) = \tilde{\mathrm{d}}(\,\tilde{v}_x\,) = 0\,, \tag{E.136}$$

so by Poincarè's Lemma, if the region $\mathcal{U}$ is star-shaped, there exists a function $H(x)$ such that $\tilde{v}_x = \tilde{\mathrm{d}}H(x)$.

Example 54. The Lie derivative of a zero-form $A(x)$ with respect to a Hamiltonian vector field $\bar{v}_x = \mathrm{d}/\mathrm{d}t|_x$, is given by:

$$\pounds_{\bar{v}_x}(A(x)) = (\,\tilde{\mathrm{d}}\,A(x)\,)(\bar{v}) = (\,\partial_\mu A(x)\,)\,\dot{x}^\mu = \frac{\mathrm{d}A(x)}{\mathrm{d}t}\,, \tag{E.137}$$

where we have used (E.100).

[12] We could just as well define the dual by putting the vector $\bar{v}$ in the *first* slot, in which case the components of the dual one-form would have opposite sign.

Example 55. Hamilton's equations are:

$$\dot{x}^\mu = f^{\mu\nu}(x)\,(\partial_\nu H(x))\,. \tag{E.138}$$

The velocity vector $\bar{v}_x$ at point $P(x)$ in $\mathcal{M}$ is defined by the convective derivative:

$$\bar{v}_x = \frac{\mathrm{d}}{\mathrm{d}t} = \dot{x}^\mu\,\bar{\partial}_\mu\,, \tag{E.139}$$

where the parameter t is time flow. So the one-form $\tilde{v}_x$ is given by:

$$\tilde{v}_x = \tilde{f}_x(\cdot,\bar{v}_x) = f_{\mu\nu}(x)\,\dot{x}^\nu\,\dot{x}^\mu = (\partial_\nu H(x))\,\tilde{\mathrm{d}}x^\nu = \tilde{\mathrm{d}}H(x)\,, \tag{E.140}$$

and is closed, $\tilde{\mathrm{d}}\tilde{v}_x = 0$. $\bar{v}_x$ is therefore a Hamiltonian vector field.

From Eq. (E.137) if $\dot{x}^\mu$ is a solution of Hamilton's equations,

$$\pounds_{\bar{v}_x}(H(x)) = \frac{\mathrm{d}H(x)}{\mathrm{d}t} = (\partial_\mu H(x))(\partial_\nu H(x))\,f^{\mu\nu}(x) = 0\,. \tag{E.141}$$

That is, $H(x)$ is a constant of the motion.

Remark 55. Using our definition of upper and lower indices, we can write the symplectic two-form in different ways:

$$\begin{aligned}\tilde{f}_x &= \frac{1}{2}\,f_{\mu\nu}(x)\,\tilde{\mathrm{d}}x^\mu\wedge\tilde{\mathrm{d}}x^\nu\\ &= \frac{1}{2}\,f_{\mu\nu}(x)\,f^{\mu'\mu}(x)\,f^{\nu'\nu}(x)\,\tilde{\mathrm{d}}x_{\mu'}\wedge\tilde{\mathrm{d}}x_{\nu'} = \frac{1}{2}\,f^{\nu'\mu'}(x)\,\tilde{\mathrm{d}}x_{\mu'}\wedge\tilde{\mathrm{d}}x_{\nu'}\\ &= -\frac{1}{2}\,f^{\mu\nu}(x)\,\tilde{\mathrm{d}}x_\mu\wedge\tilde{\mathrm{d}}x_\nu\,,\end{aligned} \tag{E.142}$$

so that the *negative* of the inverse symplectic matrix appears here with lower indices for the basis one-forms. If $\bar{a}_x$ and $\bar{b}_x$ are two vectors, then:

$$\tilde{f}_x(\bar{a}_x,\bar{b}_x) = \quad a^\mu(x)\,b^\nu(x)\,f_{\mu\nu}(x) = \quad a^\mu(x)\,b_\mu(x) \tag{E.143}$$

$$= -a_\mu(x)\,b_\nu(x)\,f^{\mu\nu}(x) = -a_\mu(x)\,b^\mu(x)\,. \tag{E.144}$$

Again, the negative sign in the last line is due to the antisymmetry of the symplectic form.

Definition 69 (Poisson bracket)**.** If $A(x)$ and $B(x)$ are functions on $\mathcal{M}$ at point P. The Poisson bracket of $A(x)$ and $B(x)$ is given by:

$$\begin{aligned}\tilde{f}(\bar{\mathrm{d}}A(x),\bar{\mathrm{d}}B(x)) &= (\partial^\mu A(x))\,f_{\mu\nu}(x)\,(\partial^\nu B(x)) = -(\partial_\mu A(x))\,f^{\mu\nu}(x)\,(\partial_\nu B(x))\\ &= -\{\,A(x),B(x)\,\}\,.\end{aligned} \tag{E.145}$$

Again, because of the antisymmetry of $f_{\mu\nu}(x)$, we have to be careful here with raising and lowering indices when passing from the first line to the second.

Theorem 66 (Jacobi's identity)**.** *The statement that $\tilde{f}$ is closed, $\tilde{\mathrm{d}}\tilde{f} = 0$, means that Jacobi's and Bianchi's identities are satisfied.*

Proof. Using the rules of exterior differentiation, we find:

$$\begin{aligned}\tilde{\mathrm{d}}\,\tilde{f} &= \frac{1}{2}\,(\tilde{\mathrm{d}}f_{\mu\nu}(x))\wedge\tilde{\mathrm{d}}x^\mu\wedge\tilde{\mathrm{d}}x^\nu\\ &= \frac{1}{2}\,(\partial_\gamma f_{\mu\nu}(x))\,\tilde{\mathrm{d}}x^\gamma\wedge\tilde{\mathrm{d}}x^\mu\wedge\tilde{\mathrm{d}}x^\nu\\ &= \frac{1}{6}\,f_{\mu\nu\lambda}(x)\,\tilde{\mathrm{d}}x^\mu\wedge\tilde{\mathrm{d}}x^\nu\wedge\tilde{\mathrm{d}}x^\gamma\,,\end{aligned} \tag{E.146}$$

where

$$f_{\mu\nu\gamma}(x) = \partial_\mu f_{\nu\lambda}(x) + \partial_\nu f_{\lambda\mu}(x) + \partial_\lambda f_{\mu\nu}(x)\,, \tag{E.147}$$

which is *odd* under interchange of all indices. So from (E.146), we find:

$$\tilde{\mathrm{d}}\,\tilde{f}(\bar{d}A, \bar{d}B, \bar{d}C) = (\,\partial_\mu A\,)\,(\,\partial_\nu B\,)\,(\,\partial_\lambda C\,)\,\{\,(\,\partial^\mu f^{\nu\lambda}\,) + (\,\partial^\nu f^{\lambda\mu}\,) + (\,\partial^\lambda f^{\mu\nu}\,)\,\}\,. \tag{E.148}$$

Resuming terms in (E.146), we can also write:

$$\tilde{\mathrm{d}}\,\tilde{f} = \tilde{\mathrm{d}}x^\mu \wedge (\,\partial_\mu\,\tilde{f}\,) = \tilde{\mathrm{d}}x^\mu \otimes (\,\partial_\mu\,\tilde{f}\,) - (\,\partial_\mu\,\tilde{f}\,) \otimes \tilde{\mathrm{d}}x^\mu\,, \tag{E.149}$$

from which we find:

$$\tilde{\mathrm{d}}\,\tilde{f}(\bar{d}A, \bar{d}B, \bar{d}C) = \{\,A, \{\,B, C\,\}\,\} + \{\,B, \{\,C, A\,\}\,\} + \{\,C, \{\,A, B\,\}\,\}\,. \tag{E.150}$$

So that using results (E.148) and (E.150), and using the fact that $\tilde{f}$ is closed, we find:

$$\begin{aligned}\tilde{\mathrm{d}}\,\tilde{f}(\bar{d}A, \bar{d}B, \bar{d}C) &= \{\,A, \{\,B, C\,\}\,\} + \{\,B, \{\,C, A\,\}\,\} + \{\,C, \{\,A, B\,\}\,\} \\ &= (\,\partial^\mu A\,)\,(\,\partial^\nu B\,)\,(\,\partial^\lambda C\,)\,\{\,(\,\partial_\mu f_{\nu\lambda}\,) + (\,\partial_\nu f_{\lambda\mu}\,) + (\,\partial_\lambda f_{\mu\nu}\,)\,\} = 0\,,\end{aligned} \tag{E.151}$$

in agreement with Theorem 62 and Bianchi's identity. □

Definition 70 (volume form)**.** The volume form $\tilde{\omega}$ for our symplectic space is given by:

$$\begin{aligned}\tilde{\Gamma} &= \underbrace{\tilde{f} \wedge \tilde{f} \wedge \cdots \wedge \tilde{f}}_{n \text{ times}} \\ &= f_{1,n+1} f_{2,n+2} \cdots f_{n,2n}\,(\tilde{\mathrm{d}}x^1 \wedge \tilde{\mathrm{d}}x^{n+1}) \wedge (\tilde{\mathrm{d}}x^2 \wedge \tilde{\mathrm{d}}x^{n+2}) \wedge \cdots \wedge (\tilde{\mathrm{d}}x^n \wedge \tilde{\mathrm{d}}x^{2n})\,.\end{aligned} \tag{E.152}$$

Remark 56. Since the space has $2n$ dimensions, $\tilde{\mathrm{d}}\,\tilde{\Gamma} = 0$. Now let us define differential *vectors* $\bar{\mathrm{d}}q = (\mathrm{d}q)\,\bar{\partial}_q$ and $\bar{\mathrm{d}}p = (\mathrm{d}p)\,\bar{\partial}_p$. Then for a *canonical* system where $f_{\mu\nu}(x)$ is independent of x and is of the block form (E.12), we have:

$$\tilde{f}(\bar{\mathrm{d}}q, \bar{\mathrm{d}}p) = \mathrm{d}q\,\mathrm{d}p\,. \tag{E.153}$$

Similarly the volume form $\tilde{\omega}$ when evaluated at vectors $\bar{\mathrm{d}}q^i$ and $\bar{\mathrm{d}}p^j$ for each value of $\mu = (i, j)$ gives, for a canonical system the usual volume element given by Eq. (E.37):

$$\mathrm{d}\Gamma = \frac{\mathrm{d}^n q\,\mathrm{d}^n p}{(2\pi\hbar)^n} = \frac{\tilde{\Gamma}(\,\bar{\mathrm{d}}q^1, \bar{\mathrm{d}}p^1, \bar{\mathrm{d}}q^2, \bar{\mathrm{d}}p^2, \ldots, \bar{\mathrm{d}}q^n, \bar{\mathrm{d}}p^n\,)}{(2\pi\hbar)^n}\,. \tag{E.154}$$

We conclude that $\tilde{\Gamma}$ is non-zero.

It is useful now to define a density of states form $\tilde{\rho}(t)$ by:

Definition 71 (density of states)**.** The density of states form $\tilde{\rho}(t)$ is defined by:

$$\tilde{\rho}(t) \equiv \rho(t)\,\tilde{\Gamma}\,, \tag{E.155}$$

where $\rho(t)$ is a function (a zero form), to be specified below. So $\tilde{\rho}(t)$ is a $2n$-form. Both $\rho(t)$ and $\tilde{\Gamma}$ depend on the coordinates x, which we have suppressed for simplicity here.

Theorem 67. *The Lie derivative of the density form $\tilde{\rho}(t)$ with respect to a Hamiltonian vector field $\bar{v}$ is given by:*

$$\pounds_{\bar{v}}(\,\tilde{\rho}(t)\,) = v^\mu\,\partial_\mu(\,\rho(t)\,)\,\tilde{\Gamma}\,, \tag{E.156}$$

from which we find that the divergence of the co-moving vector field $\rho(t)\,\bar{v}$ with respect to the volume form $\tilde{\Gamma}$ is given by:

$$\tilde{\nabla}_{\tilde{\Gamma}}(\,\rho(t)\,\bar{v}\,) = v^\mu\,\partial_\mu(\,\rho(t)\,)\,. \tag{E.157}$$

The Hamiltonian vector field $\bar{v} = \bar{\mathrm{d}}\,H(x)$ satisfies $\tilde{\mathrm{d}}\,\tilde{v} = 0$.

Proof. The Lie derivative is given by:

$$\pounds_{\bar{v}}(\,\tilde{\rho}(t)\,) = \pounds_{\bar{v}}(\,\rho(t)\,\tilde{\Gamma}\,) = \tilde{\mathrm{d}}(\,\rho(t)\,\tilde{\Gamma}(\bar{v})\,) + \tilde{\mathrm{d}}\,(\,\rho(t)\,\tilde{\Gamma}\,)(\bar{v})\,. \tag{E.158}$$

But we note that:

$$\tilde{\mathrm{d}}\,(\,\rho(t)\,\tilde{\Gamma}\,) = (\tilde{\mathrm{d}}\rho(t)) \wedge \tilde{\Gamma} = (\partial_\mu \rho(t))\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\Gamma} = 0\,, \tag{E.159}$$

since $\tilde{\Gamma}$ saturates the space. Using $\tilde{f}(\bar{v}) = \tilde{v}$, we find:

$$\tilde{\Gamma}(\bar{v}) = \tilde{v} \wedge \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} + \tilde{f} \wedge \tilde{v} \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} + \cdots + \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge \tilde{v} \tag{E.160}$$

where there are $(n-1)$ two-forms $\tilde{f}$'s. But since $\tilde{\mathrm{d}}\tilde{f} = 0$ and $\tilde{\mathrm{d}}\,\tilde{v} = 0$, we find:

$$\tilde{\mathrm{d}}(\,\rho(t)\tilde{v}\,) = \tilde{\mathrm{d}}(\,\rho(t)\,) \wedge \tilde{v} = v_\nu\,(\,\partial_\mu\,\rho(t)\,)\,(\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,)\,, \tag{E.161}$$

so that:

$$\begin{aligned}\tilde{\mathrm{d}}(\,\rho(t)\tilde{\Gamma}(\bar{v})\,) = v_\nu\,(\,\partial_\mu\,\rho(t)\,)\,\{\,(\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,) \wedge \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} \\ + \tilde{f} \wedge (\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,) \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} + \cdots + \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge (\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,)\,\}\,.\end{aligned} \tag{E.162}$$

Now using the identity:

$$\begin{aligned}f^{\nu\mu}\,\tilde{\Gamma} = (\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,) \wedge \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} \\ + \tilde{f} \wedge (\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,) \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} + \cdots + \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge (\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,)\,,\end{aligned} \tag{E.163}$$

we find:

$$\tilde{\mathrm{d}}(\,\rho(t)\,\tilde{\Gamma}(\bar{v})\,) = v_\nu\,f^{\nu\mu}\,(\,\partial_\mu\rho(t)\,)\,\tilde{\Gamma} = v^\mu\,(\,\partial_\mu\,\rho(t)\,)\,\tilde{\Gamma}\,, \tag{E.164}$$

from which the result follows. □

Exercise 86. Prove identity Eq. (E.172).

The next theorem states that the number of states in phase space, defined by an integral of the density form $\tilde{\rho}(t)$ is conserved.

Theorem 68. *If the density of states form $\tilde{\rho}(t)$ satisfies the equation:*

$$\left\{\,\frac{\partial}{\partial t} + \pounds_{\bar{v}}\,\right\}(\,\tilde{\rho}(t)\,) = 0\,, \tag{E.165}$$

then the number of states N in a region U of phase space:

$$N = \int_U \tilde{\rho}(t,\bar{v}) = \int_U \rho(t)\,\tilde{\Gamma}(\bar{v})\,, \tag{E.166}$$

is conserved, $\mathrm{d}N/\mathrm{d}t = 0$. Here U is a region of phase space where on the boundary ∂U of which $\rho(t) \to 0$.

Proof. We first note that Theorem 67 shows that Eq. (E.165) means that:

$$\frac{\partial\tilde{\rho}(t)}{\partial t} = -\tilde{\mathrm{d}}(\,\rho(t)\,\tilde{\Gamma}(\bar{v})\,)\,, \tag{E.167}$$

So that:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -\int_U \tilde{\mathrm{d}}(\,\rho(t)\,\tilde{\Gamma}(\bar{v})\,) = -\int_{\partial U} \rho(t)\,\tilde{\Gamma}(\bar{v})\big|_{\partial V} \to 0\,. \tag{E.168}$$

where we have used Stoke's theorem. □

Remark 57. Let us work out the left- and right-hand sides of Eq. (E.123). For the left-hand side, since $\tilde{f}(\bar{V}) = \tilde{V}$, we find:

$$\tilde{\omega}(\bar{V}) = \tilde{V} \wedge \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} + \tilde{f} \wedge \tilde{V} \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} + \cdots + \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge \tilde{V} \tag{E.169}$$

where there are $(n-1)$ two-forms $\tilde{f}$'s. But since $\tilde{\mathrm{d}}\tilde{f} = 0$ and

$$\tilde{\mathrm{d}}\,\tilde{V} = (\,\partial_\nu V_\mu\,)\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,, \tag{E.170}$$

we find:

$$\begin{aligned}\tilde{\mathrm{d}}(\,\tilde{\omega}(\bar{V})\,) = (\,\partial_\nu\, V_\mu\,)\,\{\,(\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,) \wedge \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} \\ + \tilde{f} \wedge (\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,) \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} + \cdots + \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge (\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,)\,\}\,.\end{aligned} \tag{E.171}$$

Now using the identity:

$$\begin{aligned}f^{\nu\mu}\,\tilde{\omega} = (\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,) \wedge \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} \\ + \tilde{f} \wedge (\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,) \wedge \tilde{f} \wedge \cdots \wedge \tilde{f} + \cdots + \tilde{f} \wedge \tilde{f} \wedge \cdots \wedge (\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,)\,,\end{aligned} \tag{E.172}$$

we find:

$$\tilde{\mathrm{d}}(\,\tilde{\omega}(\bar{V})\,) = f^{\nu\mu}\,(\,\partial_\nu\, V_\mu\,)\,\tilde{\omega} = (\,\partial^\mu\, V_\mu\,)\,\tilde{\omega}\,. \tag{E.173}$$

We also need to evaluate the right-hand side of (E.123) on the boundary ∂U. So let $\bar{s}$ be any vector tangent to ∂U and let $\tilde{n}$ be a one-form normal to ∂U so that $\tilde{n}(\bar{s}) = 0$. Then if $\tilde{S}$ is any $(N-1)$-form such that: $\tilde{\omega} = \tilde{n} \wedge \tilde{S}$, then:

$$\tilde{\omega}(\bar{V}) = \tilde{n}(\bar{V})\,\tilde{S} = (\,n_\mu V^\mu\,)\,\tilde{S}\,, \tag{E.174}$$

and so Stokes' theorem for symplectic forms becomes:

$$\int_U (\,\partial^\mu V_\mu\,)\,\tilde{\omega} = \int_{\partial U} (\,n_\mu V^\mu\,)\,\tilde{S}\,, \tag{E.175}$$

where $\tilde{S}$ is restricted to the boundary: ∂U.

E.4.2 Integral invariants

We assume that there exists a fundamental one-form $\tilde{\pi}_x \in T^*_x\mathcal{M}_S$ which describes the classical system. It is given by:

$$\tilde{\pi}_x = \pi_\mu(x)\,\tilde{\mathrm{d}}x^\mu\,. \tag{E.176}$$

The symplectic two-form is then given by:

$$\tilde{f}_x = \tilde{\mathrm{d}}\,\tilde{\pi} = (\partial_\mu\pi_\nu(x))\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu = \frac{1}{2}\,f_{\mu\nu}(x)\,\tilde{\mathrm{d}}x^\mu \wedge \tilde{\mathrm{d}}x^\nu\,, \tag{E.177}$$

where

$$f_{\mu\nu}(x) = \partial_\mu\pi_\nu(x) - \partial_\nu\pi_\mu(x) = -f_{\nu\mu}(x)\,, \tag{E.178}$$

is the antisymmetric symplectic matrix. Using Stokes' theorem, we find:

$$\int_U \tilde{f}_x = \int_U \tilde{\mathrm{d}}\,\tilde{\pi} = \int_{\partial U} \tilde{\pi}\,. \tag{E.179}$$

For the case of $n = 1$ and a cannonical coordinate system, (E.179) becomes:

$$\int_U \tilde{\mathrm{d}}p \wedge \tilde{\mathrm{d}}q = \int_{\partial U} p(q)\,\tilde{\mathrm{d}}q\,. \tag{E.180}$$

We now wish to include time as a variable. We take time to be a one dimensional base manifold, $t \in \mathbb{R}$. At each point t in this base manifold, we attach a symplectic manifold $\mathcal{M}$ of dimension $2n$ with an antisymmetric non-degenerate two-form f attached. The fiber bundle consisting of $\mathcal{M}$ plus the base manifold forms a projective manifold of dimension $2n+1$. The action one-form is given by the Poincaré-Cartan invariant:

$$\tilde{\mathrm{d}}S = \tilde{\pi}(x) - H(x)\,\tilde{\mathrm{d}}t = \pi_\mu(x)\,\tilde{\mathrm{d}}x^\mu - H(x)\,\tilde{\mathrm{d}}t \equiv L(x)\,\tilde{\mathrm{d}}t\,. \tag{E.181}$$

The integral:

$$S = \int \tilde{\mathrm{d}}S\,, \tag{E.182}$$

is called *Hilbert's invariant integral.*

E.4.3 Gauge connections

In this section, we introduce a general n-dimensional frame (a basis), and consider linear transformations of the basis set. In this section, we think of the basis set as **fields** and the basis transformations as **gauge** transformations, using the terminology of the gauge theory of fields.

Since the methods used in this section are applicable to both metric manifolds and symplectic manifolds, we develop equations for both types of manifolds. We follow the development in Göckeler and Schücker [4][p. 61] for a manifold with a metric, which is called the Einstein-Cartan general relativity theory. The book by Foster and Nightingale [**?**] is also very useful here.

Let us choose a frame[13] $\bar{b}_\mu(x) \in T_x\mathcal{M}$ at point $P(x)$ with the duals $\tilde{b}^\mu(x) \in T_x^*\mathcal{M}$. Under a basis (gauge) transformation:

$$\bar{b}'_\mu(x) = \bar{b}_\nu(x)\,[\,\gamma^{-1}(x)\,]^\nu{}_\mu\,, \qquad \tilde{b}'^{\,\mu}(x) = [\,\gamma(x)\,]^\mu{}_\nu\,\tilde{b}^\nu(x)\,, \tag{E.183}$$

where $\gamma(x) \in GL_n$. In an obvious matrix notation, we write simply:

$$\bar{b}'(x) = \bar{b}(x)\,\gamma^{-1}(x)\,, \qquad \tilde{b}'(x) = \gamma(x)\,\tilde{b}(x)\,. \tag{E.184}$$

With respect to this frame, we define a (symmetric) metric-form g_x at $P(x)$ by:

$$\mathsf{g}_x = g_{\mu\nu}(x)\,\tilde{b}^\mu(x) \otimes \tilde{b}^\nu(x)\,, \tag{E.185}$$

or a symplectic (anti-symmetric) two-form $\tilde{f}_x$ at $P(x)$ by:

$$\tilde{f}_x = \frac{1}{2}\,f_{\mu\nu}(x)\,\tilde{b}^\mu(x) \wedge \tilde{b}^\nu(x)\,, \tag{E.186}$$

depending on the type of manifold. The metric matrix $g_{\mu\nu}(x)$ then transforms according to:

$$g'_{\mu\nu}(x) = g_{\mu'\nu'}(x)\,[\,\gamma^{-1}(x)\,]^{\mu'}{}_\mu\,[\,\gamma^{-1}(x)\,]^{\nu'}{}_\nu = [\,\gamma^{-1\,T}(x)\,]_\mu{}^{\mu'}\,g_{\mu'\nu'}(x)\,[\,\gamma^{-1}(x)\,]^{\nu'}{}_\nu\,, \tag{E.187}$$

which we can write in matrix notation as:

$$g'(x) = \gamma^{-1\,T}(x)\,g(x)\,\gamma^{-1}(x)\,. \tag{E.188}$$

The symplectic matrix $f_{\mu\nu}(x)$ transforms in the same way:

$$f'(x) = \gamma^{-1\,T}(x)\,f(x)\,\gamma^{-1}(x)\,. \tag{E.189}$$

For infinitesimal transformations, we write:

$$\gamma^\mu{}_\nu(x) = \delta^\mu{}_\nu + \Delta\gamma^\mu{}_\nu(x) + \cdots\,, \tag{E.190}$$

[13] In order to avoid confusion, we set $\bar{b}_{x\,\mu} \mapsto \bar{b}_\mu(x)$ in this section.

where $\Delta\gamma^\mu{}_\nu(x) \in gl_n$ are infinitesimal. So for infinitesimal transformations, we find, in matrix notation:

$$\begin{aligned}\Delta\tilde{b}(x) &= \Delta\gamma(x)\,\tilde{b}(x)\,,\\ \Delta g(x) &= -\Delta\gamma^T(x)\,g(x) - g(x)\,\Delta\gamma(x)\,,\\ \Delta f(x) &= -\Delta\gamma^T(x)\,f(x) - f(x)\,\Delta\gamma(x)\,.\end{aligned} \tag{E.191}$$

We want to make sure that when we are done, the choice of frame is irrelevant. We seek here differential equations for $\tilde{b}^\mu(x)$ and $g_{\mu\nu}(x)$ or $f_{\mu\nu}(x)$ which are covariant under the gauge group GL_n. For this purpose, we introduce a gl_n connection, where gl_n is the Lie algebra of GL_n, and find an invariant action. Minimizing the action will lead to the equations we seek.

We first come to the problem of finding a connection. For this purpose, we seek a **covariant** exterior derivative one-form *matrix* operator $\tilde{D}^{(1)\mu}{}_\nu(x) \in gl_n$, which, when acting on the basis fields $\tilde{b}^\mu(x)$, transform homogeneously under gauge transformations. That is, we want to find a $\tilde{D}^{(1)\mu}{}_\nu(x)$ such that:

$$\tilde{D}^{(1)\prime\mu}{}_\nu(x)\,\tilde{b}'^\nu(x) = \gamma^\mu{}_\nu(x)\,\tilde{D}^{(1)\nu}{}_\lambda(x)\,\tilde{b}^\lambda(x)\,. \tag{E.192}$$

We state the result in the form of a theorem, using matrix notation:

Theorem 69 (Covariant derivative of a vector)**.** *If $\tilde{b}(x)$ transforms according to:*

$$\tilde{b}'(x) = \gamma(x)\,\tilde{b}(x)\,, \tag{E.193}$$

then the exterior covariant derivative $\tilde{D}^{(1)}(x)$ transforms according to:

$$\tilde{D}^{(1)\prime}(x)\,\tilde{b}'(x) = \tilde{D}^{(1)\prime}(x)\,\gamma(x)\,\tilde{b}(x) = \gamma(x)\,\tilde{D}^{(1)}(x)\,\tilde{b}(x)\,, \tag{E.194}$$

with $\tilde{D}^{(1)}(x)$ given by:

$$\tilde{D}^{(1)}(x) = \tilde{\mathrm{d}} + \tilde{\Gamma}\wedge\,, \tag{E.195}$$

where the one-form (matrix) connection $\tilde{\Gamma}$ transforms according to the rule:

$$\tilde{\Gamma}'(x) = \gamma(x)\,\tilde{\Gamma}(x)\,\gamma^{-1}(x) + \gamma(x)\,(\,\tilde{\mathrm{d}}\,\gamma^{-1}(x)\,)\,. \tag{E.196}$$

Proof. Using (E.193), we want to find a connection which satisfies:

$$\tilde{\mathrm{d}}\,(\,\gamma(x)\,\tilde{b}(x)\,) + \tilde{\Gamma}'(x)\,\gamma(x)\wedge\tilde{b}(x) = \gamma(x)\,(\,\tilde{\mathrm{d}}\,\tilde{b}(x)\,) + \gamma(x)\,\tilde{\Gamma}(x)\wedge\tilde{b}(x) \tag{E.197}$$

from which we find:

$$\tilde{\mathrm{d}}\,\gamma(x) + \tilde{\Gamma}'(x)\,\gamma(x) = \gamma(x)\,\tilde{\Gamma}(x)\,, \tag{E.198}$$

or:

$$\begin{aligned}\tilde{\Gamma}'(x) &= \gamma(x)\,\tilde{\Gamma}(x)\,\gamma^{-1}(x) - (\,\tilde{\mathrm{d}}\,\gamma(x)\,)\,\gamma^{-1}(x)\\ &= \gamma(x)\,\tilde{\Gamma}(x)\,\gamma^{-1}(x) + \gamma(x)\,\tilde{\mathrm{d}}\,(\,\gamma^{-1}(x)\,)\,,\end{aligned} \tag{E.199}$$

which was what we wanted to show. In the last line of (E.199) we used:

$$\tilde{\mathrm{d}}\,(\,\gamma(x)\,\gamma^{-1}(x)\,) = (\,\tilde{\mathrm{d}}\,\gamma(x)\,)\,\gamma^{-1}(x) + \gamma(x)\,(\,\tilde{\mathrm{d}}\,\gamma^{-1}(x)\,) = 0\,. \tag{E.200}$$

□

Remark 58. The transformation rule (E.196) for the connection is called an **affine** representation of the gauge group. In components, the connection can be written in terms of the basis as:

$$\tilde{\Gamma}^\mu{}_\nu(x) = \Gamma^\mu{}_{\nu\lambda}(x)\,\tilde{\mathrm{d}}x^\lambda\,, \qquad \Gamma^\mu{}_{\nu\lambda}(x) \in gl_n\,. \tag{E.201}$$

Note carefully here that λ is a *form* index whereas μ and ν are *matrix* indices which refer to gauge transformations. For infinitesimal transformations, we find from Eq. (E.196):

$$\Delta\tilde{\Gamma}(x) = -\{\,\tilde{\mathrm{d}}\,\Delta\gamma(x) + \tilde{\Gamma}(x)\,\Delta\gamma(x) - \Delta\gamma(x)\,\tilde{\Gamma}(x)\,\} = -\{\,\tilde{\mathrm{d}}\,\Delta\gamma(x) + [\,\tilde{\Gamma}(x), \Delta\gamma(x)\,]\,\}\,. \tag{E.202}$$

The connection form, or gauge fields, are considered here to be *independent* fields.

Definition 72 (Curvature and torsion)**.** The curvature $\tilde{R}(x)$ and torsion $\tilde{T}(x)$ two-forms are defined as:

$$\begin{aligned}\tilde{R}(x) &:= \tilde{D}^{(1)}\,\tilde{\Gamma}(x) = \tilde{\mathrm{d}}\,\tilde{\Gamma}(x) + \tilde{\Gamma}(x)\wedge\tilde{\Gamma}(x)\,,\\ \tilde{T}(x) &:= \tilde{D}^{(1)}\,\tilde{b}(x) \;= \tilde{\mathrm{d}}\,\tilde{b}(x) \;+ \tilde{\Gamma}(x)\wedge\tilde{b}(x)\,.\end{aligned} \tag{E.203}$$

Remark 59. In the definition of the curvature, we have encountered the wedge product of two connection one-forms. This term can be written in a number of ways:

$$\begin{aligned}[\,\tilde{\Gamma}(x)\wedge\tilde{\Gamma}(x)\,]^{\mu}{}_{\nu} &= \tilde{\Gamma}^{\mu}{}_{\lambda}(x)\wedge\tilde{\Gamma}^{\lambda}{}_{\nu}(x)\\ &= \Gamma^{\mu}{}_{\lambda\sigma}(x)\,\Gamma^{\lambda}{}_{\nu\sigma'}(x)\,\tilde{b}^{\sigma}(x)\wedge\tilde{b}^{\sigma'}(x)\\ &= \frac{1}{2}\left\{\,\Gamma^{\mu}{}_{\lambda\sigma}(x)\,\Gamma^{\lambda}{}_{\nu\sigma'}(x) - \Gamma^{\mu}{}_{\lambda\sigma'}(x)\,\Gamma^{\lambda}{}_{\nu\sigma}(x)\,\right\}\tilde{b}^{\sigma}(x)\wedge\tilde{b}^{\sigma'}(x)\\ &= \frac{1}{2}\,[\,\Gamma_{\sigma}(x),\Gamma_{\sigma'}(x)\,]^{\mu}{}_{\nu}\,\tilde{b}^{\sigma}(x)\wedge\tilde{b}^{\sigma'}(x)\,.\end{aligned} \tag{E.204}$$

Since the commutator in this expression is a *matrix* product, it does not in general vanish. Another way of saying this is that the connection one-form $\tilde{\Gamma}\in gl_n$ and obeys a Lie algegra. So it is useful to define a commutation relation for matrices of one-forms belonging to a Lie algebra by:

Definition 73 (Wedge commutator)*.* The wedge commutator of two one-forms $\tilde{A},\tilde{B}\in gl_4$ is defined by:

$$[\,\tilde{A},\tilde{B}\,]_{\wedge} := [\,A_{\mu},B_{\nu}\,]\,\tilde{b}^{\mu}(x)\wedge\tilde{b}^{\nu}(x) = \tilde{A}\wedge\tilde{B} - \tilde{B}\wedge\tilde{A}\,. \tag{E.205}$$

In particular if $\mathtt{T}_a$ is a basis of gl_n and $\tilde{A}^i$ are p-forms and $\tilde{B}^j$ q-forms, the wedge commutator becomes:

$$\tilde{A} = \tilde{A}^i\,\mathtt{T}_i\,,\qquad \tilde{B} = \tilde{B}^j\,\mathtt{T}_j\,, \tag{E.206}$$

so that

$$[\,\tilde{A},\tilde{B}\,]_{\wedge} = \tilde{A}^i\wedge\tilde{B}^j\,[\,\mathtt{T}_i,\mathtt{T}_j\,]\,. \tag{E.207}$$

This definition means that:

$$[\,\tilde{A},\tilde{B}\,]_{\wedge} = -(-)^{pq}\,[\,\tilde{B},\tilde{A}\,]_{\wedge}\,, \tag{E.208}$$

and in particular $[\,\tilde{A},\tilde{A}\,]_{\wedge}\neq 0$ for odd-forms.

Using this definition, the curvature can be written as:

$$\tilde{R}(x) = \tilde{\mathrm{d}}\,\tilde{\Gamma}(x) + \frac{1}{2}\,[\,\tilde{\Gamma},\tilde{\Gamma}\,]_{\wedge}\,. \tag{E.209}$$

We next find the transformation properties of the curvature and torsion. This is stated in the following theorem:

Theorem 70. *The curvature and torsion transform homogeneously:*

$$\tilde{R}'(x) = \gamma(x)\,\tilde{R}(x)\,\gamma^{-1}(x)\,,\qquad \textit{and}\qquad \tilde{T}'(x) = \gamma(x)\,\tilde{T}(x)\,. \tag{E.210}$$

For infinitesimal transformations, we have:

$$\Delta\tilde{R}(x) = [\,\Delta\gamma(x),\tilde{R}(x)\,]\,,\qquad \textit{and}\qquad \Delta\tilde{T}(x) = \Delta\gamma(x)\,\tilde{T}(x)\,. \tag{E.211}$$

Proof. This is easily established using definitions (E.203), and is left as an exercise for the reader. □

Remark 60. The Bianchi identities for the curvature and torsion are:

$$\tilde{D}^{(1)}\,\tilde{T}(x) = \tilde{D}^{(1)}\,\tilde{D}^{(1)}\,\tilde{b}(x) = \tilde{R}(x)\wedge\tilde{b}(x)\,, \tag{E.212}$$

which can easily be established using the definitions. The Bianchi identity for the curvature requires a definition of the covariant derivative of a tensor, which we state in the following theorem.

Theorem 71 (Covariant derivative of a tensor)**.** *If $\tilde{R}(x)$ transforms according to:*

$$\tilde{R}'(x) = \gamma(x)\,\tilde{R}(x)\,\gamma^{-1}(x)\,, \tag{E.213}$$

then the exterior covariant derivative $\tilde{D}^{(2)}(x)$ transforms according to:

$$\tilde{D}^{(2)\,\prime}(x)\,\tilde{R}'(x) = \tilde{D}^{(2)\,\prime}(x)\,\gamma(x)\,\tilde{R}(x)\,\gamma^{-1}(x) = \gamma(x)\,\tilde{D}^{(2)}(x)\,\tilde{R}(x)\,\gamma^{-1}(x)\,, \tag{E.214}$$

with $\tilde{D}^{(2)}(x)\,\tilde{R}(x)$ given by:

$$\tilde{D}^{(2)}(x)\,\tilde{R}(x) = \tilde{\mathrm{d}}\,\tilde{R}(x) + [\,\tilde{\Gamma}(x),\tilde{R}(x)\,]_{\wedge}\,. \tag{E.215}$$

with the connection form $\tilde{\Gamma}(x)$ transforming according to the rule given in Eq. (E.196)*.*

Proof. We have, in a short-hand notation:

$$\begin{aligned}\tilde{D}^{(2)\,\prime}\,\tilde{R}' &= \tilde{D}^{(2)\,\prime}\,\gamma\,\tilde{R}\,\gamma^{-1}\\ &= \tilde{\mathrm{d}}\,(\,\gamma\,\tilde{R}\,\gamma^{-1}\,) + [\,\tilde{\Gamma}',\gamma\,\tilde{R}\,\gamma^{-1}\,]_{\wedge}\\ &=\end{aligned} \tag{E.216}$$

□

References

[1] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics* (Addison-Wesley, Reading, MA, 2002), third edition.

[2] M. Tabor, *Chaos and Integrability in Nonlinear Dyanmics, an Introduction* (John Wiley & Sons, New York, NY, 1989).

[3] H. Flanders, *Differential forms with applications to the physical sciences* (1989). Original version published in 1963 by Academic Press, New York.

[4] M. Göckeler and T. Schücker, *Differential geometry,gGauge theories, and gravity*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 1987).

Appendix F

Statistical mechanics review

In this appendix we review statistical mechanics as applied to quantum mechanical systems. We start with definitions of the canonical and grand canonical thermal density matrices and their relationship to thermodynamic quantities. We discuss classical perturbative methods for computing statistical density matrices for interacting systems, illustrated by the anharmonic oscillator. Finally, we discuss the Martin-Siggia-Rose (MSR) [1, 2] method of finding a generating function for classical perturbation theory.

F.1 Thermal ensembles

We first study a classical system of N particles with dynamics governed by a Hamiltonian H.

F.2 Grand canonical ensemble

The thermal density matrix $\hat{\rho}$ for a grand canonical ensemble is generated by writing the entropy (S), energy (E), and particle number (N) in terms of a quantum density operator $\hat{\rho}$ and quantum operators $\hat{H}$ and $\hat{N}$.[1]

$$S = -k_{\rm B}\,{\rm Tr}[\,\hat{\rho}\,\ln\hat{\rho}\,]\,, \qquad E = {\rm Tr}[\,\hat{\rho}\,\hat{H}\,]\,, \qquad N = {\rm Tr}[\,\hat{\rho}\,\hat{N}\,]\,, \qquad 1 = {\rm Tr}[\,\hat{\rho}\,]\,. \tag{F.1}$$

Here $k_{\rm B}$ is Boltzmann's constant. The best choice of the density $\hat{\rho}$ which minimizes the entropy, such that the energy, particle number, and normalization is conserved is given by:

$$\hat{\rho} = \frac{1}{Z}\,e^{-\beta\,(\hat{H}-\mu\hat{N})}\,, \tag{F.2}$$

where β, μ, and Z are Lagrange multipliers. This density matrix is *not* idempotent, and so it cannot be represented by a projection operator for a single quantum state $|\,\psi\,\rangle\langle\,\psi\,|$. Instead it is a mixture of many quantum states.

The partition function $Z(\beta,\mu,V)$ is given by requiring the thermal density matrix to be normalized to one:

$$Z(\beta,\mu,V) \equiv e^{-W(\beta,\mu,V)} = {\rm Tr}[\,e^{-\beta(\hat{H}-\mu\hat{N})}\,]\,. \tag{F.3}$$

Here we have define a "connnected generator" $W(\beta,\mu,V)$, which will be useful below. V is the volume of the system, and enters through the trace definition. Derivatives of $W(\beta,\mu,V)$ with respect to β and μ are given:

$$-\,\frac{1}{Z}\left[\,\frac{\partial Z(\beta,\mu,V)}{\partial\beta}\,\right]_{\mu,V} = \left[\,\frac{\partial W(\beta,\mu,V)}{\partial\beta}\,\right]_{\mu,V} = {\rm Tr}[\,\hat{\rho}\,(\hat{H}-\mu\hat{N})\,] = E-\mu N\,, \tag{F.4}$$

[1] In this section, hatted quantities are quantum operators.

and

$$\frac{1}{Z}\left[\frac{\partial Z(\beta,\mu,V)}{\partial\mu}\right]_{\beta,V} = -\left[\frac{\partial W(\beta,\mu,V)}{\partial\mu}\right]_{\beta,V} = \beta\,\mathrm{Tr}[\,\hat{\rho}\,\hat{N}\,] = \beta\,N\,. \tag{F.5}$$

The values of β and μ are fixed by these two equations. The entropy is now given by:

$$S(E,N,V)/k_{\rm B} = -\mathrm{Tr}[\,\hat{\rho}\,\ln\hat{\rho}\,] = \beta\,(E-\mu N) + \ln Z(\beta,\mu,V) = \beta\,(E-\mu N) - W(\beta,\mu,V)\,. \tag{F.6}$$

But Eq. (F.6) is a Legendre transformation to the entropy S which is now a function of E, N, and V as a result of Eqs. (F.4) and (F.5). So we now find:

$$\left[\frac{\partial S(E,N,V)}{\partial E}\right]_{N,V} = k_{\rm B}\,\beta\,, \qquad \left[\frac{\partial S(E,N,V)}{\partial N}\right]_{E,V} = -k_{\rm B}\,\beta\,\mu\,, \tag{F.7}$$

$$\left[\frac{\partial S(E,N,V)}{\partial V}\right]_{E,N} = -k_{\rm B}\left[\frac{\partial W(\beta,\mu,V)}{\partial V}\right]_{\beta,\mu}\,. \tag{F.8}$$

The first and second laws of thermodynamics tells us that:

$$T\,\mathrm{d}S(E,N,V) = \mathrm{d}E - \mu\,\mathrm{d}N + p\,\mathrm{d}V\,, \tag{F.9}$$

which means that that:

$$\left[\frac{\partial S(E,N,V)}{\partial E}\right]_{N,V} = \frac{1}{T}\,, \qquad \left[\frac{\partial S(E,N,V)}{\partial N}\right]_{E,V} = -\frac{\mu}{T}\,, \qquad \left[\frac{\partial S(E,N,V)}{\partial V}\right]_{E} = \frac{p}{T}\,. \tag{F.10}$$

But from Eqs. (F.7) and (F.8), we see that $\beta = 1/(k_{\rm B}T)$ is the inverse temperature times $k_{\rm B}$ and μ is called the chemical potential. The pressure is found from the equation:

$$p = T\left[\frac{\partial S(E,N,V)}{\partial V}\right]_{E,N} = -\frac{1}{\beta}\left[\frac{\partial W(\beta,\mu,V)}{\partial V}\right]_{\beta,\mu} = -\left[\frac{\partial\Omega(\beta,\mu,V)}{\partial V}\right]_{\beta,\mu}\,, \tag{F.11}$$

where we have defined a thermodynamic potential $\Omega(\beta,\mu,V) = W(\beta,\mu,V)/\beta$, so that now $Z(\beta,\mu,V)$ is related to the connected generator $\Omega(\beta,\mu,V)$ by:

$$Z(\beta,\mu,V) = e^{-\beta\,\Omega(\beta,\mu,V)}\,. \tag{F.12}$$

In terms of the thermodynamic potential $\Omega(\beta,\mu,V)$ the number of particles and energy is given by:

$$N = -\left[\frac{\partial\Omega(\beta,\mu,V)}{\partial\mu}\right]_{\beta,V}\,, \qquad E = \left[\frac{\partial\,[\,\beta\,\Omega(\beta,\mu,V)\,]}{\partial\beta}\right]_{\mu,V} + \mu\,N\,. \tag{F.13}$$

This completes the identification of the thermodynamic variables. So what we have learned here is that if we can find the generating function $Z(\beta,\mu,V)$, we can identify β with the inverse temperature times $k_{\rm B}$ and μ with the chemical potential and thus the number of particles N. The pressure is then calculated from Eq. (F.11), which enables us to find the equation of state, $p = p(T,N,V)$ for the system. We will illustrate this with some examples below.

F.2.1 The canonical ensemble

For systems of particles which do not conserve particle number, such as photons or pions, we use a canonical ensemble. This is the same as the grand canonical ensemble with $\mu = 0$. That is the best choice of the density $\hat{\rho}$ function which minimizes the entropy, such that the energy and normalization is conserved is given by:

$$\hat{\rho}(q,p) = \frac{e^{-\beta H(q,p)}}{Z(\beta)}\,, \qquad Z(\beta) = \mathrm{Tr}[e^{-\beta H}]\,. \tag{F.14}$$

F.3 Some examples

We start with a classical example of a N-free particles with the Hamiltonian:

$$H(p) = \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m} \,. \tag{F.15}$$

In the classical case, the trace for calculation of the grand canonical ensemble is defined by:

$$Z(\beta,\mu) = \iint_{-\infty}^{+\infty} \prod_{i=1}^{N} \frac{\mathrm{d}^3 q_i \,\mathrm{d}^3 p_i}{(2\pi\hbar)^{3N}} \exp\Big\{ -\beta\Big[\sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m} - \mu\, N \Big] \Big\} \tag{F.16}$$

Integration gives:

$$Z(\beta,\mu) = \Big[\frac{V}{(2\pi\hbar)^3} \Big]^N \Big[\frac{2m\pi}{\beta} \Big]^{3N/2} e^{\beta\,\mu\, N} = \exp\Big\{ \beta\mu N + N \,\ln\big[\, V/(2\pi\hbar)^3 \,\big] + \frac{3N}{2} \,\ln\big[\, 2m\pi/\beta \,\big] \Big\} \,. \tag{F.17}$$

So

$$\Omega(\beta,\mu,V) = -\mu N - \frac{N}{\beta} \,\ln\big[\, V/(2\pi\hbar)^3 \,\big] - \frac{3N}{2\beta} \,\ln\big[\, 2m\pi/\beta \,\big] \,. \tag{F.18}$$

The pressure is then:

$$p = -\Big[\frac{\partial\Omega(\beta,\mu,V)}{\partial V} \Big]_{\beta,\mu} = \frac{N}{\beta V} \,. \tag{F.19}$$

Setting $\beta = 1/(k_{\mathrm{B}}T)$ and putting $R = N_{\mathrm{A}}/k_{\mathrm{B}}$ and $n = N/N_A$, where n is the number of moles and N_{A} is Avogardo's number, we find the ideal gas law: $pV = nRT$. As expected, we find:

$$\begin{aligned} N &= -\Big[\frac{\partial\Omega(\beta,\mu,V)}{\partial\mu} \Big]_{\beta,V} \,, \\ E &= \Big[\frac{\partial\,[\,\beta\,\Omega(\beta,\mu,V)\,]}{\partial\beta} \Big]_{\mu,V} + \mu\, N = \frac{3}{2}\,\frac{1}{\beta} = \frac{3}{2}\,k_{\mathrm{B}}\,T = \frac{3}{2}\,n\,RT \,. \end{aligned} \tag{F.20}$$

The distribution function is the classical Boltzmann distribution:

$$\rho(q,p) = \frac{1}{Z} \exp\Big\{ -\beta\Big[\sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m} - \mu\, N \Big] \Big\} \,. \tag{F.21}$$

Note that nothing here depends on $\hbar$, which was used only to make Z dimensionless.

F.4 Martin-Siggia-Rose formalism

The purpose of the Martin-Siggia-Rose (MSR) development is to find a generating function which can be used to develop diagrammatic rules for classical perturbation expansions [1, 2] . We consider the case for a anharmonic classical oscillator with a Hamiltonian of the form:

$$H(q,p) = \frac{1}{2}\,[\,p^2 + \mu^2\,q^2\,] + \frac{\lambda}{8}\,q^4 \,. \tag{F.22}$$

Here the classical canonical variables are q and p. The classical equations of motion, Eq. (**??**), are:

$$\begin{aligned} \dot{q} &= \{\, q, H \,\} = p \,, \\ \dot{p} &= \{\, p, H \,\} = -\mu^2\, q - \lambda\, q^3/2 \,. \end{aligned} \tag{F.23}$$

The *classical* Poisson bracket is defined by:

$$\{ A(q,p), B(q,p) \} = \frac{\partial A}{\partial q}\frac{\partial B}{\partial p} - \frac{\partial B}{\partial q}\frac{\partial A}{\partial p} \,. \tag{F.24}$$

q and q satisfy the Poisson bracket relations:

$$\{ q, p \} = 1 \,, \qquad \{ q, q \} = \{ p, p \} = 0 \,. \tag{F.25}$$

So the classical equations of motion can be written in the form:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} p \\ -\mu^2\, q - \lambda\, q^3/2 \end{pmatrix} \,. \tag{F.26}$$

In the MSR formalism, we define quantum *operators* Q and P which satisfy quantum *comutation* relations with the *classical* canonical variables q and p such that:

$$\begin{aligned} [\, q, P \,] = [\, Q, p \,] &= i\hbar \,, \\ [\, q, q \,] = [\, P, P \,] = [\, Q, Q \,] = [\, p, p \,] &= 0 \,, \\ [\, q, Q \,] = [\, q, p \,] = [\, P, Q \,] = [\, p, P \,] &= 0 \,, \end{aligned} \tag{F.27}$$

Evidently, a representation of P and Q acting on *functions* of q and p is given by the differential operators:

$$P \to \frac{\hbar}{i}\frac{\partial}{\partial q} \,, \qquad Q \to -\frac{\hbar}{i}\frac{\partial}{\partial p} \,. \tag{F.28}$$

This satisfies all the commutation relations (F.27). Next, we define a "non-hermitian Hamiltonian" H_{MSR} by:

$$H_{\mathrm{MSR}}[q,p;P,Q] = P\, p + Q\, \{\, \mu^2 q + \lambda\, q^3/2 \,\} \,. \tag{F.29}$$

H_{MSR} is an *operator* with scalar coefficients. Using the *quantum* Heisenberg equations of motion,

$$\begin{aligned} \dot{q} &= [\, q, H_{\mathrm{MSR}} \,]/i\hbar \;\; = \{\, q, H \,\} = p \,, \\ \dot{p} &= [\, p, H_{\mathrm{MSR}} \,]/i\hbar \;\; = \{\, p, H \,\} = -\mu^2 q - \lambda\, q^3/2 \,, \\ \dot{P} &= [\, P, \hat{H}_{\mathrm{MSR}} \,]/i\hbar = -(\, \mu^2 + 3\lambda\, q^2/2 \,)\, Q \,, \\ \dot{Q} &= [\, Q, H_{\mathrm{MSR}} \,]/i\hbar = P \,, \end{aligned} \tag{F.30}$$

we recover the original classical equations of motion for q and p, but find two additional equations of motion for the quantum operators Q and P. The two second order equations of motion are obtained from these by differentiation. We find:

$$\begin{aligned} \ddot{q} + \mu^2\, q + \lambda\, q^3/2 &= 0 \,, \\ \ddot{Q} + \mu^2\, Q + 3\lambda\, q^2\, Q/2 &= 0 \,. \end{aligned} \tag{F.31}$$

It is now useful to change variables in the following way. We let:

$$\begin{aligned} Q_\pm = q \pm \frac{Q}{2} &= \quad q \mp \frac{\hbar}{2i}\frac{\partial}{\partial p} \,, \\ P_\pm = \frac{P}{2} \pm p &= \pm p + \frac{\hbar}{2i}\frac{\partial}{\partial q} \,. \end{aligned} \tag{F.32}$$

Solving (F.32) in reverse, we find:

$$\begin{aligned} q &= (\, Q_+ + Q_- \,)/2 \,, \qquad Q = Q_+ - Q_- \,, \\ p &= (\, P_+ - P_- \,)/2 \,, \qquad P = P_+ + P_- \,. \end{aligned} \tag{F.33}$$

So we have the relations:

$$\begin{aligned}
\frac{1}{2}\,[\,P_+^2 - P_-^2\,] &= P\,p \\
\frac{1}{2}\,[\,Q_+^2 - Q_-^2\,] &= q\,Q \\
\frac{1}{8}\,[\,Q_+^4 - Q_-^4\,] &= q^3\,Q/2 + q\,Q^3/8\,,
\end{aligned} \tag{F.34}$$

$Q_\pm$ and $P_\pm$ are sums of functions and operators, so they obey commutation relations as well as Poisson bracket relations. These new operators $Q_\pm$ and $P_\pm$ satisfy the commutation relations:

$$\begin{aligned}
[\,Q_+, P_+\,] = [\,Q_-, P_-\,] &= i\hbar\,, \\
[\,Q_+, Q_+\,] = [\,Q_-, Q_-\,] = [\,Q_+, Q_-\,] &= 0\,, \\
[\,P_+, P_+\,] = [\,P_-, P_-\,] = [\,P_+, P_-\,] &= 0\,, \\
[\,Q_+, Q_-\,] = [\,Q_-, P_+\,] &= 0\,.
\end{aligned} \tag{F.35}$$

and the Poisson bracket relations:

$$\{\,Q_+, P_\pm\,\} = \{\,Q_-, P_\pm\,\} = \pm 1\,. \tag{F.36}$$

Now as we have seen, the quantum mechanical closed-time-path Hamiltonian for this problem is given by the difference of two identical Hamiltonians with varibles (Q_+, P_+) and (Q_-, P_-) respectively:

$$H_{\rm CTP} = H[Q_+, P_+] - H[Q_-, P_-] = \frac{1}{2}\,[\,P_+^2 - P_-^2\,] + \frac{\mu^2}{2}\,[\,Q_+^2 - Q_-^2\,] + \frac{\lambda}{8}\,[\,Q_+^4 - Q_-^4\,]\,, \tag{F.37}$$

which can be written as:

$$H_{\rm CTP}[q, p; Q, P] = p\,P + \mu^2\,q\,Q + \lambda\,[\,q^3\,Q/2 + q\,P^3/8\,]\,. \tag{F.38}$$

Thus, except for the factor $q\,Q^3/8$ in the last term, the CTP Hamiltonian, Eq. (F.38), is the same as the MSR Hamiltonian, Eq. (F.29). So if the CTP Hamiltonian is to reduce to the MSR Hamiltonian, we must retain Q and P as quantum variables, and set q and p to be classical variables. Then, in the classical limit, $\hbar \to 0$, according to the differential representations (F.28), the *first* term in the last equation in (F.34) is of order $\hbar$, whereas the *second* term is of order $\hbar^3$, and is thus to be neglected. In which case

$$H_{\rm CTP} \to H_{\rm MSR} \tag{F.39}$$

in the classical limit. H still has a term linear in $\hbar$, but no classical equations depend on $\hbar$, as we have shown in the first of Eqs. (F.31).

We can now reverse the whole arguement, and develop a rule for moving from quantum mechanics to classical physics to obtain a generating function. Starting with the quantum CTP Lagrangian or Hamiltonian with $(Q_+, P_+; Q_-, P_-)$ variables, replace them with the set $(q, p; Q, P)$ set of variables, with q and p classical. Retain the first order in $\hbar$ reduction of the quantum variables, Q and P, and the resulting Lagrangian or Hamiltonian will generate the *classical* equations of motion, as well as quantum ones for Q and P.

We note from Eq. (F.30) that $\dot q = p$ and $\dot Q = P$, so if we define the Lagrangian by

$$\begin{aligned}
L[q, Q; \dot q, \dot Q] &= \dot q\,P + \dot Q\,p - H[\hat q, Q; p, \hat P]\,, \\
&= \dot Q\,\dot q - Q\,\{\,\mu^2 q + \lambda\,q^3/2\,\}\,,
\end{aligned} \tag{F.40}$$

it gives the equations of motion:

$$\begin{aligned}
\frac{\rm d}{{\rm d}t}\,\frac{\partial L}{\partial \dot Q} - \frac{\partial L}{\partial Q} = 0\,, \quad &\to \qquad \ddot q + \mu^2\,q + \lambda\,q^3/2 = 0\,, \\
\frac{\rm d}{{\rm d}t}\,\frac{\partial L}{\partial \dot q} - \frac{\partial L}{\partial q} = 0\,, \quad &\to \quad \ddot Q + \mu^2\,Q + \lambda\,3\,q^2\,Q/2 = 0\,,
\end{aligned} \tag{F.41}$$

which agree with our previous result, Eqs. (F.31). Introducing current terms into the Lagrangian, we can write Eq. (F.40) as:

$$\begin{aligned} L'[q,Q;\dot{q},\dot{Q}] &= \dot{Q}\,\dot{q} - Q\,\{\,\mu^2 q + \lambda\,q^3/2\,\} + q\,J + Q\,j\,, \\ &= -Q\left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \mu^2\right\} q - \frac{\lambda}{2}\,Q\,q^3 + q\,J + Q\,j\,. \end{aligned} \tag{F.42}$$

in which case, the equations of motion become:

$$\begin{aligned} \ddot{q} + \mu^2\,q + \lambda\,q^3/2 &= j\,, \\ \ddot{Q} + \mu^2\,Q + \lambda\,3\,q^2\,Q/2 &= J\,, \end{aligned} \tag{F.43}$$

It is useful to introduce two component vectors $\mathcal{Q}^a = (q, Q)$ and $\mathcal{J}^a = (j, J)$, and a metric g^{ab}, given by:

$$g^{ab} = g_{ab} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \tag{F.44}$$

which raise and lower indicies, so that $\mathcal{Q}_a = (Q, q)$ and $\mathcal{J}_a = (J, j)$. Then, from Eq. (F.42), the action can be written in the compact form:

$$\begin{aligned} S[\mathcal{Q}] = &-\frac{1}{2}\iint \mathrm{d}t\,\mathrm{d}t'\;\mathcal{Q}^a(t)\,G^{-1}_{0\,ab}(t,t')\,\mathcal{Q}^b(t') \\ &- \int \mathrm{d}t\left\{\frac{1}{4}\,\gamma_{abcd}\,\mathcal{Q}^a(t)\,\mathcal{Q}^b(t)\,\mathcal{Q}^c(t)\,\mathcal{Q}^d(t) - \mathcal{J}_a(t)\,\mathcal{Q}^a(t)\right\}. \end{aligned} \tag{F.45}$$

where $G^{-1}_{0\,ab}(t,t')$ and γ_{abcd} are given by,

$$G^{-1}_{0\,ab}(t,t') = \left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \mu^2\right\} g_{ab}\,\delta(t-t')\,. \tag{F.46}$$

$$\gamma_{\hat{q}QQQ} = \gamma_{Q\hat{q}QQ} = \gamma_{QQ\hat{q}Q} = \gamma_{QQQ\hat{q}} = \frac{\lambda}{2}\,, \tag{F.47}$$

all other γ's vanish. The γ's are fully symmetric with respect to permutations of the indices:

$$\gamma_{abcd} = \gamma_{bacd} = \gamma_{cbad} = \gamma_{dbca} = \gamma_{acbd} = \cdots \tag{F.48}$$

In this new notation, the equations of motion (F.43) are given by the single equation:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \mu^2\right\} \mathcal{Q}_a(t) + \gamma_{abcd}\,\mathcal{Q}^b(t)\,\mathcal{Q}^c(t)\,\mathcal{Q}^d(t) = \mathcal{J}_a(t)\,. \tag{F.49}$$

We can also define the vector $\mathcal{P}^a = \dot{\mathcal{Q}}^a = (p, \hat{P})$. Then,

$$[\,\mathcal{Q}^a(t), \mathcal{P}^b(t)\,] = i\hbar\,g^{ab}\,, \qquad [\,\mathcal{Q}^a(t), \mathcal{Q}^b(t)\,] = [\,\mathcal{P}^a(t), \mathcal{P}^b(t)\,] = 0\,. \tag{F.50}$$

Keep in mind that we are doing *classical* physics here! What we are trying to do is to bring out the similarities and differences between classical and quantum perturbative calculations.

F.4.1 Classical statistical averages

We now come to the tricky point of defining a statistical average of our classical and (quantum) operator quantities. Using a canonical ensemble, we define the statistical average of a function of Q and p by Eq. (**??**), which here reads:

$$\langle\,A(q,p)\,\rangle = \mathrm{Tr}[\,A(q,p)\,] = \frac{1}{Z}\iint_{-\infty}^{+\infty}\frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,A(q,p)\,e^{-\beta H(q,p)}\,. \tag{F.51}$$

We have supressed the time dependence of $A(q,p)$. It is useful to note that $H(q,p)$ and the phase space element $\mathrm{d}q\,\mathrm{d}p$ are invariant under a time translation generated by $H(q,p)$.

For the operators Q and P, a different strategy for computing an average is required. We define these as right-acting differential operators, given by Eq. (F.28), and define a generalized ensemble average by:

$$\langle\, A(q,Q;p,P)\,\rangle = \mathrm{Tr}[\,A(q,Q;p,P)\,] = \frac{1}{Z}\iint_{-\infty}^{+\infty}\frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,A(q,Q;p,P)\,e^{-\beta H(q,p)}\,, \tag{F.52}$$

where now $A(q,Q;p,P)$ is a differential operator. For example, we find:

$$\langle\, Q(t)\,\rangle = -\frac{\hbar}{iZ}\iint_{-\infty}^{+\infty}\frac{\mathrm{d}q(t)\,\mathrm{d}p(t)}{2\pi\hbar}\,\frac{\partial}{\partial p(t)}\left\{\,e^{-\beta H(q(t),p(t))}\,\right\} = 0\,, \tag{F.53}$$

by integration by parts. We also find that:

$$\langle\, Q(t)\,q(t')\,\rangle = -\frac{\hbar}{iZ}\iint_{-\infty}^{+\infty}\frac{\mathrm{d}q(t)\,\mathrm{d}p(t)}{2\pi\hbar}\,\frac{\partial}{\partial p(t)}\left\{\,q(t')\,e^{-\beta H(q(t),p(t))}\,\right\} = 0\,, \tag{F.54}$$

Here we used the fact that the phase space $\mathrm{d}q(t)\,\mathrm{d}p(t)$ and Hamiltonian $H(q(t),p(t))$ are invariant under time translations. In a similar way, we find that

$$\langle\, P(t)\,p(t')\,\rangle = 0\,. \tag{F.55}$$

In fact the extended definition of the ensemble average, given in Eq. (F.52), shows that the average vanishes if any *operator* stands first on the left. This is crucial for the definition of the tri-diagonal form of the Green functions, as we will see below. If the operator stands last on the right, the situation is altogether different. For example, consider:

$$\begin{aligned}
\langle\, q(t)\,Q(t')\,\rangle &= -\frac{\hbar}{iZ}\iint_{-\infty}^{+\infty}\frac{\mathrm{d}q(t')\,\mathrm{d}p(t')}{2\pi\hbar}\,q(t)\,\frac{\partial}{\partial p(t')}\,e^{-\beta H(q(t'),p(t'))}\\
&= \frac{\hbar}{iZ}\iint_{-\infty}^{+\infty}\frac{\mathrm{d}q(t')\,\mathrm{d}p(t')}{2\pi\hbar}\,e^{-\beta H(q(t'),p(t'))}\,\frac{\partial q(t)}{\partial p(t')}\\
&= -\frac{\hbar}{iZ}\iint_{-\infty}^{+\infty}\frac{\mathrm{d}q(t')\,\mathrm{d}p(t')}{2\pi\hbar}\,e^{-\beta H(q(t'),p(t'))}\left\{\frac{\partial q(t)}{\partial q(t')}\,\frac{\partial q(t')}{\partial p(t')} - \frac{\partial q(t)}{\partial p(t')}\,\frac{\partial q(t')}{\partial q(t')}\right\}\\
&= i\hbar\,\langle\,\{\,q(t),q(t')\,\}\,\rangle = i\hbar\,\sigma(t,t')\,,
\end{aligned} \tag{F.56}$$

since

$$\frac{\partial q(t')}{\partial q(t')} = 1\,, \quad \text{and} \quad \frac{\partial q(t')}{\partial p(t')} = 0\,.$$

The Poisson bracket in Eq. (F.56) can be evaluated at any time, in particular at $t=0$. From (F.54) and (F.56), we obtain:

$$\sigma(t,t') = \langle\,\{\,q(t),q(t')\,\}\,\rangle = \langle\,[\,q(t),Q(t')\,]\,\rangle/i\hbar\,. \tag{F.57}$$

Here, the first expression is classical and the second is an operator statement.

In a similar way, we find:

$$\begin{aligned}
\langle\, p(t)\,\hat{P}(t')\,\rangle &= \frac{\hbar}{iZ}\iint_{-\infty}^{+\infty}\frac{\mathrm{d}Q(t')\,\mathrm{d}p(t')}{2\pi\hbar}\,p(t)\,\frac{\partial}{\partial Q(t')}\,e^{-\beta H(Q(t'),p(t'))}\\
&= -\frac{\hbar}{iZ}\iint_{-\infty}^{+\infty}\frac{\mathrm{d}Q(t')\,\mathrm{d}p(t')}{2\pi\hbar}\,e^{-\beta H(Q(t'),p(t'))}\,\frac{\partial p(t)}{\partial Q(t')}\\
&= -\frac{\hbar}{iZ}\iint_{-\infty}^{+\infty}\frac{\mathrm{d}Q(t')\,\mathrm{d}p(t')}{2\pi\hbar}\,e^{-\beta H(Q(t'),p(t'))}\left\{\frac{\partial p(t)}{\partial Q(t')}\,\frac{\partial p(t')}{\partial p(t')} - \frac{\partial p(t)}{\partial p(t')}\,\frac{\partial p(t')}{\partial Q(t')}\right\}\\
&= i\hbar\,\langle\,\{\,p(t),p(t')\,\}\,\rangle = i\hbar\,\langle\,\{\,\dot{Q}(t),\dot{Q}(t')\,\}\,\rangle = i\hbar\,\frac{\partial^2\sigma(t,t')}{\partial t\,\partial t'}\,.
\end{aligned} \tag{F.58}$$

So we can write:

$$\frac{\partial^2\sigma(t,t')}{\partial t\,\partial t'} = \langle\,\{\,p(t),p(t')\,\}\,\rangle = \langle\,[\,p(t),\hat{P}(t')\,]\,\rangle/i\hbar\,. \tag{F.59}$$

Time- and antitime-ordered products are defined by:

$$\begin{aligned}\mathcal{T}\{\,Q(t)\,\hat{q}(t')\,\} &= Q(t)\,\hat{q}(t')\,\Theta(t-t') + \hat{q}(t')\,Q(t)\,\Theta(t'-t)\,,\\ \mathcal{T}^*\{\,Q(t)\,\hat{q}(t')\,\} &= Q(t)\,\hat{q}(t')\,\Theta(t'-t) + \hat{q}(t')\,Q(t)\,\Theta(t-t') = \mathcal{T}\{\,\hat{q}(t')\,Q(t)\,\}\,.\end{aligned} \tag{F.60}$$

So again, using properties (F.54) and (F.56), we find:

$$\begin{aligned}-\langle\,\mathcal{T}\{\,Q(t)\,\hat{q}(t')\,\}\,\rangle/i\hbar &= -\langle\,\{\,Q(t),Q(t')\,\}\,\rangle\,\Theta(t-t') = -\sigma(t,t')\,\Theta(t-t') = G_{\rm R}(t,t')/i\,,\\ +\langle\,\mathcal{T}^*\{\,Q(t)\,\hat{q}(t')\,\}\,\rangle/i\hbar &= +\langle\,\{\,Q(t),Q(t')\,\}\,\rangle\,\Theta(t'-t) = +\sigma(t,t')\,\Theta(t'-t) = G_{\rm A}(t,t')/i\,.\end{aligned}$$

Except for the factor of $\hbar$, this is the same as we found before. Note that $G_{\rm A}(t,t') = G_{\rm R}(t',t)$, so that we only need to consider time-ordered operators.

Using the two-component notation, the ensemble average is defined by:

$$\langle\,A(\mathcal{Q},\mathcal{P})\,\rangle = {\rm Tr}[\,A(\mathcal{Q},\mathcal{P})\,] = \frac{1}{Z}\iint_{-\infty}^{+\infty}\frac{{\rm d}Q\,{\rm d}p}{2\pi\hbar}\,A(\mathcal{Q},\mathcal{P})\,e^{-\beta H(Q,p)}\,. \tag{F.61}$$

F.4.2 Generating functions

Green functions

Using the two-component notation, the generating functional for extended ensemble averages, given in the last section, of time-orded products of $\mathcal{Q}_a(t_a)$ is defined by:

$$Z[\mathcal{J}] = e^{iW[\mathcal{J}]/\hbar} = \left\langle\,\mathcal{T}\left\{\,\exp\left\{\frac{i}{\hbar}\int_{-\infty}^{+\infty}{\rm d}t\,\mathcal{Q}_a(t)\mathcal{J}^a(t)\right\}\right\}\right\rangle\,. \tag{F.62}$$

Then the time-ordered product is given by:

$$\langle\,\mathcal{T}\{\,\mathcal{Q}_a(t_a)\mathcal{Q}_b(t_b)\mathcal{Q}_c(t_c)\,\ldots\}\,\rangle = \left(\frac{\hbar}{i}\right)^n\frac{1}{Z}\left[\,\frac{\delta^n Z[\mathcal{J}]}{\delta\mathcal{J}^a(t_a)\,\delta\mathcal{J}^b(t_b)\,\delta\mathcal{J}^c(t_c)\,\cdots}\,\right]_{\mathcal{J}=0}\,. \tag{F.63}$$

We define connected Green functions by the relation:

$$W_{abc\cdots}(t_a,t_b,t_c,\ldots) = \left(\frac{\hbar}{i}\right)^{n-1}\left[\,\frac{\delta^n W[\mathcal{J}]}{\delta\mathcal{J}^a(t_a)\,\delta\mathcal{J}^b(t_b)\,\delta\mathcal{J}^c(t_c)\,\cdots}\,\right]_{\mathcal{J}=0}\,. \tag{F.64}$$

The order n of the Green functions are given by the number of indices. So then first order Green functions are average values:

$$\langle\,\mathcal{Q}_a(t)\,\rangle = \left(\frac{\hbar}{i}\right)\frac{1}{Z}\left[\,\frac{\delta Z[\mathcal{J}]}{\delta\mathcal{J}^a(t)}\right]_{\mathcal{J}=0} = \left[\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}^a(t)}\right]_{\mathcal{J}=0} = W_a(t)\,, \tag{F.65}$$

and second order ones are correlation coefficients and the usual Green functions:

$$\begin{aligned}\langle\,\mathcal{T}\{\,\mathcal{Q}_a(t_a)\mathcal{Q}_b(t_b)\,\}\,\rangle &= \left(\frac{\hbar}{i}\right)^2\frac{1}{Z}\left[\,\frac{\delta^2 Z[\mathcal{J}]}{\delta\mathcal{J}^a(t_a)\,\delta\mathcal{J}^b(t_b)}\right]_{\mathcal{J}=0}\\ &= \langle\,\mathcal{Q}_a(t_a)\,\rangle\langle\,\mathcal{Q}_b(t_b)\,\rangle + \frac{\hbar}{i}\left[\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}^a(t_a)\,\delta\mathcal{J}^b(t_b)}\right]_{\mathcal{J}=0}\\ &= \langle\,\mathcal{Q}_a(t_a)\,\rangle\langle\,\mathcal{Q}_b(t_b)\,\rangle + W_{ab}(t_a,t_b)\,.\end{aligned} \tag{F.66}$$

So

$$W_{ab}(t_a,t_b) = \langle\, \mathcal{T}\{\, \mathcal{Q}_a(t_a)\mathcal{Q}_b(t_b)\,\}\,\rangle - \langle\, \mathcal{Q}_a(t_a)\,\rangle\langle\, \mathcal{Q}_b(t_b)\,\rangle\,. \tag{F.67}$$

Explicitly, we find for the upper component Green functions:

$$\begin{aligned} W^{QQ}(t,t') &= \langle\, \mathcal{T}\{\, Q(t)Q(t')\,\}\,\rangle - \langle\, Q(t)\,\rangle\langle\, Q(t')\,\rangle \\ &= \langle\, Q(t)Q(t')\,\rangle - \langle\, Q(t)\,\rangle\langle\, Q(t')\,\rangle = F(t,t')\,, \end{aligned} \tag{F.68}$$

$$\begin{aligned} W^{Q\hat{q}}(t,t') &= \langle\, \mathcal{T}\{\, Q(t)\hat{q}(t')\,\}\,\rangle - \langle\, Q(t)\,\rangle\langle\, \hat{q}(t')\,\rangle \\ &= \langle\, Q(t)\hat{q}(t')\,\rangle\,\Theta(t-t') \\ &= i\hbar\,\langle\, \{\, Q(t), Q(t')\,\}\,\rangle\,\Theta(t-t') \\ &= i\hbar\,\sigma(t,t')\,\Theta(t-t') = -i\hbar\,G_{\rm R}(t,t')\,. \end{aligned} \tag{F.69}$$

$$\begin{aligned} W^{\hat{q}Q}(t,t') &= \langle\, \mathcal{T}\{\, \hat{q}(t)Q(t')\,\}\,\rangle - \langle\, \hat{q}(t)\,\rangle\langle\, Q(t')\,\rangle \\ &= \langle\, Q(t')\hat{q}(t)\,\rangle\,\Theta(t'-t) \\ &= i\hbar\,\langle\, \{\, Q(t'), Q(t)\,\}\,\rangle\,\Theta(t'-t) \\ &= i\hbar\,\sigma(t',t)\,\Theta(t'-t) = -i\hbar\,G_{\rm A}(t,t')\,, \end{aligned} \tag{F.70}$$

$$W^{\hat{q}\hat{q}}(t,t') = \langle\, \mathcal{T}\{\, \hat{q}(t)\hat{q}(t')\,\}\,\rangle - \langle\, \hat{q}(t)\,\rangle\langle\, \hat{q}(t')\,\rangle = 0\,. \tag{F.71}$$

So

$$\begin{aligned} iW^{ab}(t,t') &= \begin{pmatrix} iF(t,t') & \hbar\,G_{\rm R}(t,t') \\ \hbar\,G_{\rm A}(t,t') & 0 \end{pmatrix}\,, \\ iW_{ab}(t,t') &= \begin{pmatrix} 0 & \hbar\,G_{\rm A}(t,t') \\ \hbar\,G_{\rm R}(t,t') & iF(t,t') \end{pmatrix}\,. \end{aligned} \tag{F.72}$$

The *order* of differentiation doesn't matter since $\mathcal{J}_a(t)$ is considered to be a classical commuting variable. So, for example, $W_{ab}(t,t') = W_{ba}(t',t)$, as can be seen explicitly above.

We will also need third order Green functions. These are given by:

$$\begin{aligned} &\langle\, \mathcal{T}\{\, \mathcal{Q}_a(t_a)\mathcal{Q}_b(t_b)\mathcal{Q}_c(t_c)\,\}\,\rangle = \\ &\qquad \left(\frac{\hbar}{i}\right)^3 \frac{1}{Z}\left[\,\frac{\delta^3 Z[\mathcal{J}]}{\delta\mathcal{J}^a(t_a)\,\delta\mathcal{J}^b(t_b)\,\mathcal{J}^c(t_c)}\right]_{\mathcal{J}=0} = \langle\, \mathcal{Q}_a(t_a)\,\rangle\langle\, \mathcal{Q}_b(t_b)\,\rangle\langle\, \mathcal{Q}_c(t_c)\,\rangle \\ &+ \frac{\hbar}{i}\left[\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}^a(t_a)}\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}^b(t_b)\,\delta\mathcal{J}^c(t_c)} + \frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}^b(t_b)}\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}^c(t_c)\,\delta\mathcal{J}^a(t_a)} + \frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}^c(t_c)}\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}^a(t_a)\,\delta\mathcal{J}^b(t_b)}\right]_{\mathcal{J}=0} \\ &\qquad + \left(\frac{\hbar}{i}\right)^2\left[\,\frac{\delta^3 W[\mathcal{J}]}{\delta\mathcal{J}^a(t_a)\,\delta\mathcal{J}^b(t_b)\,\delta\mathcal{J}^c(t_c)}\right]_{\mathcal{J}=0} \\ &= \langle\, \mathcal{Q}_a(t_a)\,\rangle\langle\, \mathcal{Q}_b(t_b)\,\rangle\langle\, \mathcal{Q}_c(t_c)\,\rangle + \langle\, \mathcal{Q}_a(t_a)\,\rangle\, W_{bc}(t_b,t_c) + \langle\, \mathcal{Q}_b(t_b)\,\rangle\, W_{ca}(t_c,t_a) \\ &\qquad + \langle\, \mathcal{Q}_c(t_c)\,\rangle\, W_{ab}(t_a,t_b) + W_{abc}(t_a,t_b,t_c)\,. \end{aligned}$$

So

$$\begin{aligned} W_{abc}(t_a,t_b,t_c) = \langle\, \mathcal{T}\{\, \mathcal{Q}_a(t_a)\mathcal{Q}_b(t_b)\mathcal{Q}_c(t_c)\,\}\,\rangle &- \langle\, \mathcal{Q}_a(t_a)\,\rangle\, W_{bc}(t_b,t_c) - \langle\, \mathcal{Q}_b(t_b)\,\rangle\, W_{ca}(t_c,t_a) \\ &- \langle\, \mathcal{Q}_c(t_c)\,\rangle\, W_{ab}(t_a,t_b) - \langle\, \mathcal{Q}_a(t_a)\,\rangle\langle\, \mathcal{Q}_b(t_b)\,\rangle\langle\, \mathcal{Q}_c(t_c)\,\rangle\,. \end{aligned} \tag{F.73}$$

Vertex functions

The inverse (vertex) functions are obtained by a Legendre transform. We define $\Gamma[\mathcal{Q}]$ by:

$$\Gamma[\mathcal{Q}] = \int \mathrm{d}t\, \mathcal{J}_a(t)\, \mathcal{Q}^a(t) - W[\mathcal{J}]\,. \tag{F.74}$$

Here we have set the *ensemble* average $\mathcal{Q}_a(t) = \langle\, \mathcal{Q}_a(t)\,\rangle = W_a^{(1)}(t)$, which is a classical commuting variable[2]. So then

$$\frac{\delta W[\mathcal{J}]}{\delta \mathcal{J}_a(t_a)} = \mathcal{Q}^a(t_a)\,, \qquad \frac{\delta \Gamma[\mathcal{Q}]}{\delta \mathcal{Q}^a(t_a)} = \mathcal{J}_a(t_a)\,. \tag{F.75}$$

In general, we define:

$$\Gamma_{abc\cdots}(t_a,t_b,t_c,\ldots) = \left(\frac{i}{\hbar}\right)^{n-1}\left[\,\frac{\delta^n \Gamma[\mathcal{Q}]}{\delta\mathcal{Q}^a(t_a)\,\delta\mathcal{Q}^b(t_b)\,\delta\mathcal{Q}^c(t_c)\,\cdots}\,\right]_{\mathcal{Q}=0}\,. \tag{F.76}$$

In order for this to make sense, we must be able to solve (F.75) for $\mathcal{Q}^a(t_a)$ as a functional of $\mathcal{J}_a(t_a)$. We assume that this is always possible to do.

Differentiating expressions (F.75), we find:

$$\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_b(t_b)\,\delta\mathcal{J}_a(t_a)} = \frac{\delta\mathcal{Q}^a(t_a)}{\delta\mathcal{J}_b(t_b)}\,, \qquad \frac{\delta^2\Gamma[\mathcal{Q}]}{\delta\mathcal{Q}^b(t_b)\,\delta\mathcal{Q}^a(t_a)} = \frac{\delta\mathcal{J}_a(t_a)}{\delta\mathcal{Q}^b(t_b)}\,. \tag{F.77}$$

But, by the chain rule,

$$\int \mathrm{d}t_b\,\frac{\delta\mathcal{J}_b(t_b)}{\delta\mathcal{Q}^a(t_a)}\,\frac{\delta\mathcal{Q}^c(t_c)}{\delta\mathcal{J}_b(t_b)} = \delta_a{}^c\,\delta(t_a-t_c)\,, \tag{F.78}$$

we find:

$$\int \mathrm{d}t_b\,\frac{\delta^2\Gamma[\mathcal{Q}]}{\delta\mathcal{Q}^a(t_a)\,\delta\mathcal{Q}^b(t_b)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_b(t_b)\,\delta\mathcal{J}_c(t_c)} = \delta_a{}^c\,\delta(t_a-t_c)\,. \tag{F.79}$$

We can write this as:

$$\int \mathrm{d}t_b\,\Gamma_{ab}[\mathcal{Q}](t_a,t_b)\,W^{bc}[\mathcal{J}](t_b,t_c) = \delta_a{}^c\,\delta(t_a-t_c)\,, \tag{F.80}$$

so that $\Gamma_{ab}[\mathcal{Q}](t_{a,b})$ is the inverse of $W^{bc}[\mathcal{J}](t_b,t_c)$. Note that $g_a{}^c = \delta_a{}^c$. In a similar way, we can show that

$$\int \mathrm{d}t_b\,W^{ab}[\mathcal{J}](t_a,t_b)\,\Gamma_{bc}[\mathcal{Q}](t_b,t_c) = \delta^a{}_c\,\delta(t-t'')\,. \tag{F.81}$$

Differentiating (F.80) with respect to $\mathcal{J}_d(t_d)$ gives:

$$\int \mathrm{d}t_b\left\{\frac{\delta\Gamma_{ab}[\mathcal{Q}](t_a,t_b)}{\delta\mathcal{J}_d(t_d)}\,W^{bc}[\mathcal{J}](t_b,t_c) + \Gamma_{ab}[\mathcal{Q}](t_a,t_b)\,\frac{\delta W^{bc}[\mathcal{J}](t_b,t_c)}{\delta\mathcal{J}_d(t_d)}\right\} = 0\,. \tag{F.82}$$

Now using:

$$\begin{aligned}
\frac{\delta W^{bc}[\mathcal{J}](t_b,t_c)}{\delta\mathcal{J}_d(t_d)} &= W^{dbc}[\mathcal{J}](t_d,t_b,t_c)\,,\\
\frac{\delta\Gamma_{ab}[\mathcal{Q}](t_a,t_b)}{\delta\mathcal{J}_d(t_d)} &= \int \mathrm{d}t_e\,\frac{\delta Q^e(t_e)}{\delta\mathcal{J}_d(t_d)}\,\frac{\delta\Gamma_{ab}[\mathcal{Q}](t_a,t_b)}{\delta Q^e(t_e)}\\
&= \int \mathrm{d}t_e\,W^{de}[\mathcal{J}](t_d,t_e)\,\Gamma_{eab}[\mathcal{Q}](t_e,t_a,t_b)\,.
\end{aligned} \tag{F.83}$$

So (F.82) becomes:

$$\int \mathrm{d}t_b\,\Gamma_{ab}[\mathcal{Q}](t_a,t_b)\,W^{dbc}[\mathcal{J}](t_d,t_b,t_c) = \\ -\iint \mathrm{d}t_b\,\mathrm{d}t_e\,W^{de}[\mathcal{J}](t_d,t_e)\,\Gamma_{eab}[\mathcal{Q}](t_e,t_a,t_b)\,W^{bc}[\mathcal{J}](t_b,t_c) \tag{F.84}$$

[2]There is a confusion of symbols here. In order to follow the usual notation, we have used $\mathcal{Q}_a(t)$ both as the ensemble average and an operator. One can distinguish the difference between the two by the context in which is it used.

Now multiplying through by $W^{fa}(t_f,t_a)$, integrating over t_a and using (F.81) gives:

$$W^{abc}[\mathcal{J}](t_a,t_b,t_c) = -\iiint \mathrm{d}t_{a'}\,\mathrm{d}t_{b'}\,\mathrm{d}t_{c'} \times W^{aa'}[\mathcal{J}](t_a,t_{a'})\,W^{bb'}[\mathcal{J}](t_b,t_{b'})\,\Gamma_{a'b'c'}[\mathcal{Q}](t_{a'},t_{b'},t_{c'})\,W^{c'c}[\mathcal{J}](t_{c'},t_c)\,. \quad \text{(F.85)}$$

Differentiating this expression with respect to $\mathcal{J}_d(t_d)$, and using the chain rule again, gives:

$$\begin{aligned} W^{dabc}[\mathcal{J}](t_d,t_a,t_b,t_c) = &-\iiint \mathrm{d}t_{a'}\,\mathrm{d}t_{b'}\,\mathrm{d}t_{c'} \\ &\times\Big\{W^{daa'}[\mathcal{J}](t_d,t_a,t_{a'})\,W^{bb'}[\mathcal{J}](t_b,t_{b'})\,\Gamma_{a'b'c'}[\mathcal{Q}](t_{a'},t_{b'},t_{c'})\,W^{c'c}[\mathcal{J}](t_{c'},t_c) \\ &+W^{aa'}[\mathcal{J}](t_a,t_{a'})\,W^{dbb'}[\mathcal{J}](t_d,t_b,t_{b'})\,\Gamma_{a'b'c'}[\mathcal{Q}](t_{a'},t_{b'},t_{c'})\,W^{c'c}[\mathcal{J}](t_{c'},t_c) \\ &+W^{aa'}[\mathcal{J}](t_a,t_{a'})\,W^{bb'}[\mathcal{J}](t_b,t_{b'})\,\Gamma_{a'b'c'}[\mathcal{Q}](t_{a'},t_{b'},t_{c'})\,W^{dc'c}[\mathcal{J}](t_d,t_{c'},t_c)\Big\} \\ &-\iiint\!\!\int \mathrm{d}t_{a'}\,\mathrm{d}t_{b'}\,\mathrm{d}t_{c'}\,\mathrm{d}t_{d'}\,\Big\{W^{dd'}[\mathcal{J}](t_d,t_{d'})\,W^{aa'}[\mathcal{J}](t_a,t_{a'})\,W^{bb'}[\mathcal{J}](t_b,t_{b'}) \\ &\times\Gamma_{d'a'b'c'}[\mathcal{Q}](t_{d'},t_{a'},t_{b'},t_{c'})\,W^{c'c}[\mathcal{J}](t_{c'},t_c)\Big\}\,. \end{aligned} \quad \text{(F.86)}$$

F.4.3 Schwinger-Dyson equations

Using the two-component notation, the equations of motion, Eq. (F.49), can be written:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2}+\mu^2\right\} g_{ab}\,\mathcal{Q}^b(t) + \gamma_{abcd}\,\mathcal{Q}^b(t)\,\mathcal{Q}^c(t)\,\mathcal{Q}^d(t) = \mathcal{J}_a(t)\,. \quad \text{(F.87)}$$

The ensemble average of this equation is given by:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2}+\mu^2\right\} g_{ab}\,\langle\,\mathcal{Q}^b(t)\,\rangle + \gamma_{abcd}\,\langle\,\mathcal{Q}^b(t)\,\mathcal{Q}^c(t)\,\mathcal{Q}^d(t)\,\rangle = \mathcal{J}_a(t)\,. \quad \text{(F.88)}$$

Now

$$\langle\,\mathcal{Q}^b(t)\,\rangle = \frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_b(t)}\,, \quad \text{(F.89)}$$

and

$$\begin{aligned} \langle\,\mathcal{Q}^b(t)\,\mathcal{Q}^c(t)\,\mathcal{Q}^d(t)\,\rangle = \langle\,\mathcal{T}\{\,\mathcal{Q}^b(t)\,\mathcal{Q}^c(t)\,\mathcal{Q}^d(t)\,\}\,\rangle &= \frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_b(t)}\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_c(t)}\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_d(t)} \\ &+\frac{\hbar}{i}\left[\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_b(t)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_c(t)\,\delta\mathcal{J}_d(t)}+\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_c(t)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_d(t)\,\delta\mathcal{J}_b(t)}+\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_d(t)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_b(t)\,\delta\mathcal{J}_c(t)}\right] \\ &+\left(\frac{\hbar}{i}\right)^2\left[\frac{\delta^3 W[\mathcal{J}]}{\delta\mathcal{J}_b(t)\,\delta\mathcal{J}_c(t)\,\delta\mathcal{J}_d(t)}\right] \end{aligned}$$

It is only when these quantities are evaluated at $\mathcal{J}=0$ that they become generalized Green functions. Eq. (F.88) is to be regarded as a functional differential equation for $W[\mathcal{J}]$:

$$\begin{aligned} \left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2}+\mu^2\right\} g_{ab}\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_b(t)} + \gamma_{abcd}\Bigg\{&\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_b(t)}\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_c(t)}\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_d(t)} \\ &+\frac{\hbar}{i}\left[\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_b(t)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_c(t)\,\delta\mathcal{J}_d(t)}+\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_c(t)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_d(t)\,\delta\mathcal{J}_b(t)}+\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_d(t)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_b(t)\,\delta\mathcal{J}_c(t)}\right] \\ &+\left(\frac{\hbar}{i}\right)^2\left[\frac{\delta^3 W[\mathcal{J}]}{\delta\mathcal{J}_b(t)\,\delta\mathcal{J}_c(t)\,\delta\mathcal{J}_d(t)}\right]\Bigg\} = \mathcal{J}_a(t)\,. \end{aligned} \quad \text{(F.90)}$$

Differentiating (F.90) with respect to $\mathcal{J}_e(t')$ gives:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2}+\mu^2\right\} g_{ab}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_b(t)}+\gamma_{abcd}\left\{\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_b(t)}\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_c(t)}\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_d(t)}\right.$$
$$+\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_b(t)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_c(t)}\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_d(t)}+\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_b(t)}\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_c(t)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_d(t)}$$
$$+\frac{\hbar}{i}\left[\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_b(t)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_c(t)\,\delta\mathcal{J}_d(t)}+\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_c(t)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_d(t)\,\delta\mathcal{J}_b(t)}+\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_d(t)}\,\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_b(t)\,\delta\mathcal{J}_c(t)}\right]$$
$$+\frac{\hbar}{i}\left[\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_b(t)}\,\frac{\delta^3 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_c(t)\,\delta\mathcal{J}_d(t)}+\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_c(t)}\,\frac{\delta^3 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_d(t)\,\delta\mathcal{J}_b(t)}\right.$$
$$\left.\left.+\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_d(t)}\,\frac{\delta^3 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_b(t)\,\delta\mathcal{J}_c(t)}\right]+\left(\frac{\hbar}{i}\right)^2\left[\frac{\delta^4 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_b(t)\,\delta\mathcal{J}_c(t)\,\delta\mathcal{J}_d(t)}\right]\right\}=\delta_{ae}\,\delta(t-t')\,. \quad \text{(F.91)}$$

We define $\Gamma_{0\,ab}(t,t')$ by:

$$\Gamma_{0\,ab}(t,t')=\left(\frac{i}{\hbar}\right)\left[\frac{\mathrm{d}^2}{\mathrm{d}t^2}+\mu^2\right] g_{ab}\,\delta(t-t')\,. \quad \text{(F.92)}$$

Collecting terms using the symmetry of γ_{abcd}, Eq. (F.91) becomes:

$$\int \mathrm{d}t''\,\Gamma_{0\,ab}(t,t'')\left(\frac{\hbar}{i}\right)\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_b(t'')}+\gamma_{abcd}\left(\frac{i}{\hbar}\right)\left\{3\left[\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_c(t)}\,\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_d(t)}\right.\right.$$
$$\left.+\left(\frac{\hbar}{i}\right)\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_c(t)\,\delta\mathcal{J}_d(t)}\right]\left(\frac{\hbar}{i}\right)\frac{\delta^2 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_b(t)}+3\left(\frac{\hbar}{i}\right)^2\left[\frac{\delta W[\mathcal{J}]}{\delta\mathcal{J}_b(t)}\,\frac{\delta^3 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_c(t)\,\delta\mathcal{J}_d(t)}\right]$$
$$\left.\left.+\left(\frac{\hbar}{i}\right)^3\left[\frac{\delta^4 W[\mathcal{J}]}{\delta\mathcal{J}_e(t')\,\delta\mathcal{J}_b(t)\,\delta\mathcal{J}_c(t)\,\delta\mathcal{J}_d(t)}\right]\right\}=\delta_a{}^e\,\delta(t-t')\,. \quad \text{(F.93)}$$

So we obtain:

$$\int \mathrm{d}t''\,\Gamma_{0\,ab}(t,t'')\,W^{be}[\mathcal{J}](t'',t')$$
$$+\gamma_{abcd}\left(\frac{i}{\hbar}\right)\left\{\,3\left\{\,W^c[\mathcal{J}](t)\,W^d[\mathcal{J}](t)+W^{cd}[\mathcal{J}](t,t)\,\right\}W^{be}[\mathcal{J}](t,t')\right.$$
$$\left.+3\,W^b[\mathcal{J}](t)\,W^{cde}[\mathcal{J}](t,t,t')+W^{bcde}[\mathcal{J}](t,t,t,t')\,\right\}=\delta_a{}^e\,\delta(t-t')\,. \quad \text{(F.94)}$$

Evaluating this at $\mathcal{J}=0$, gives an equation connecting the second order Green functions to higher ones. Multiplying Eq. (F.94) by $\Gamma_{ee'''}[\mathcal{Q}](t',t''')$, suming over e and integrating over t' gives:

$$\Gamma_{ae}[\mathcal{Q}](t,t'')=\Gamma_{0\,ae}(t,t'')$$
$$+\gamma_{aecd}\left(\frac{i}{\hbar}\right)3\left\{\,\mathcal{Q}^c(t)\,\mathcal{Q}^d(t)+W^{cd}[\mathcal{J}](t,t)\,\right\}\delta(t-t')$$
$$+\gamma_{abcd}\left(\frac{i}{\hbar}\right)\int \mathrm{d}t'\left\{\,3\,\mathcal{Q}^b(t)\,W^{cde'}[\mathcal{J}](t,t,t')+W^{bcde'}[\mathcal{J}](t,t,t,t')\,\right\}\Gamma_{e'e}[\mathcal{Q}](t',t'') \quad \text{(F.95)}$$

Inserting (F.85) and (F.86) into (F.95) gives:

$$\begin{aligned}
\Gamma_{ae}[\mathcal{Q}](t,t'') &= \Gamma_{0\,ae}(t,t'') \\
&\quad + \gamma_{aecd}\left(\frac{i}{\hbar}\right) 3\left\{\,\mathcal{Q}^c(t)\,\mathcal{Q}^d(t) + W^{cd}[\mathcal{J}](t,t)\,\right\}\delta(t-t') \\
&\quad - \gamma_{abcd}\left(\frac{i}{\hbar}\right)\left\{ 3\,\mathcal{Q}^b(t)\iint \mathrm{d}t_{a'}\,\mathrm{d}t_{b'}\,W^{ca'}[\mathcal{J}](t,t_{a'})\,W^{db'}[\mathcal{J}](t,t_{b'})\,\Gamma_{a'b'e}[\mathcal{Q}](t_{a'},t_{b'},t'')\right. \\
&\quad\quad - \iiint \mathrm{d}t_{a'}\,\mathrm{d}t_{b'}\,\mathrm{d}t_{c'} \\
&\quad\quad \times\Big\{W^{daa'}[\mathcal{J}](t_d,t_a,t_{a'})\,W^{bb'}[\mathcal{J}](t_b,t_{b'})\,\Gamma_{a'b'c'}[\mathcal{Q}](t_{a'},t_{b'},t_{c'})\,W^{c'c}[\mathcal{J}](t_{c'},t_c) \\
&\quad\quad + W^{aa'}[\mathcal{J}](t_a,t_{a'})\,W^{dbb'}[\mathcal{J}](t_d,t_b,t_{b'})\,\Gamma_{a'b'c'}[\mathcal{Q}](t_{a'},t_{b'},t_{c'})\,W^{c'c}[\mathcal{J}](t_{c'},t_c) \\
&\quad\quad + W^{aa'}[\mathcal{J}](t_a,t_{a'})\,W^{bb'}[\mathcal{J}](t_b,t_{b'})\,\Gamma_{a'b'c'}[\mathcal{Q}](t_{a'},t_{b'},t_{c'})\,W^{dc'c}[\mathcal{J}](t_d,t_{c'},t_c)\Big\} \\
&\quad\quad - \iiiint \mathrm{d}t_{a'}\,\mathrm{d}t_{b'}\,\mathrm{d}t_{c'}\,\mathrm{d}t_{d'}\Big\{W^{dd'}[\mathcal{J}](t_d,t_{d'})\,W^{aa'}[\mathcal{J}](t_a,t_{a'})\,W^{bb'}[\mathcal{J}](t_b,t_{b'}) \\
&\quad\quad\quad \times\Gamma_{d'a'b'c'}[\mathcal{Q}](t_{d'},t_{a'},t_{b'},t_{c'})\,W^{c'c}[\mathcal{J}](t_{c'},t_c)\Big\}\Big\}\Gamma_{e'e}[\mathcal{Q}](t',t'')\,. \\
&\quad + \gamma_{abcd}\left(\frac{i}{\hbar}\right)\int \mathrm{d}t'\left\{\,3\,\mathcal{Q}^b(t)\,W^{cde'}[\mathcal{J}](t,t,t') + W^{bcde'}[\mathcal{J}](t,t,t,t')\,\right\}\Gamma_{e'e}[\mathcal{Q}](t',t'')
\end{aligned} \tag{F.96}$$

F.5 The classical anharmonic oscillator

In classical physics, the anharmonic oscillator can be scaled in the following way. We first let $t' = \mu t$. Then the Lagrangian (**??**) becomes:

$$L(q,\dot{q}) = \frac{\mu^2}{2}\left[\left(\frac{\mathrm{d}q}{\mathrm{d}t'}\right)^2 - q^2\right] - \frac{\lambda}{8}\,q^4\,. \tag{F.97}$$

So if we now scale q by $q = \alpha q'$, we find

$$L(q',\dot{q}') = \frac{\mu^2\,\alpha^2}{2}\left[\left(\frac{\mathrm{d}q'}{\mathrm{d}t'}\right)^2 - q'^{\,2}\right] - \frac{\lambda\alpha^4}{8}\,q'^{\,4}\,. \tag{F.98}$$

So the requirement of scaling is that

$$\mu^2\,\alpha^2 = \kappa\,\lambda\alpha^4 \qquad \text{or,} \qquad \alpha^2 = \frac{\mu^2}{\kappa\lambda}\,, \tag{F.99}$$

with κ arbitrary. Then (F.98) becomes:

$$L(q',\dot{q}') = \frac{\mu^4}{\kappa\lambda}\left\{\frac{1}{2}\left[\left(\frac{\mathrm{d}q'}{\mathrm{d}t'}\right)^2 - q'^{\,2}\right] - \frac{1}{8\kappa}\,q'^{\,4}\right\} = \frac{\mu^4}{\kappa\lambda}L'(q',\dot{q}')\,. \tag{F.100}$$

So we have:

$$t = t'/\mu\,, \qquad q = \frac{\mu}{\sqrt{\kappa\lambda}}\,q'\,, \qquad L = \frac{\mu^4}{\kappa\lambda}\,L'\,, \tag{F.101}$$

with

$$p = \frac{\partial L}{\partial\dot{q}} = \dot{q} = \frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\mu^2}{\sqrt{\kappa\lambda}}\,\frac{\mathrm{d}q'}{\mathrm{d}t'} = \frac{\mu^2}{\sqrt{\kappa\lambda}}\,\frac{\partial L'}{\partial\dot{q}'} = \frac{\mu^2}{\sqrt{\kappa\lambda}}\,p'\,, \tag{F.102}$$

where the new Lagrangian is given by:

$$L'(q',\dot{q}') = \frac{1}{2}\left[\left(\frac{\mathrm{d}q'}{\mathrm{d}t'}\right)^2 - q'^2\right] - \frac{1}{8\kappa}\,q'^4\,, \tag{F.103}$$

If we choose $\kappa = 1$, then the new Lagrangian is the same as the old one with $\mu = \lambda = 1$. Another choice is $\kappa = 3$, in which case the new Lagrangian is the same as the old one if we set $\mu = 1$ and $\lambda = 1/\kappa = 1/3$. The reason for using this choice for κ seems to be that sometimes the self-interaction term in the Lagrangian is written as $\lambda/4!$ rather that $\lambda/8$, which is what we use here. This second choice is given by setting $\lambda = 1$ rather that $1/3$ in the original Lagrangian. Of course, it doesn't matter, and for our purposes, it is simpler to just take $\kappa = 1$.

So using this convention, the equation of motion for the scaled variables are:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}t'^2} + 1\right] q'(t') = -\frac{1}{4}\,q'^3(t')\,. \tag{F.104}$$

The Hamiltonian scales like the Lagrangian, so if we require $\beta H = \beta' H'$, we have:

$$H'(q',p') = \frac{1}{2}\,[p'^2 + q'^2] + \frac{1}{8}\,q'^4\,, \qquad \beta = \frac{\lambda}{\mu^4}\,\beta'\,. \tag{F.105}$$

The phase space volume scales according to:

$$\mathrm{d}q\,\mathrm{d}p = \frac{\mu^3}{\lambda}\,\mathrm{d}q'\,\mathrm{d}p'\,, \tag{F.106}$$

so the partition function scales according to:

$$Z(\beta) = \iint_{-\infty}^{+\infty} \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,e^{-\beta H(q,p)} = \frac{\mu^3}{\hbar\lambda}\iint_{-\infty}^{+\infty} \frac{\mathrm{d}q'\,\mathrm{d}p'}{2\pi}\,e^{-\beta' H'(q',p')} = \frac{\mu^3}{\hbar\lambda}\,I(\beta')\,. \tag{F.107}$$

Here $I(\beta')$ is a universal function of $\beta' = \beta\mu^4/\lambda$.

F.5.1 The partition function for the anharmonic oscillator

We first look at the partition function. We expand Z in powers of the interaction λ. From (**??**), we find:

$$\begin{aligned} Z(\beta) &= \iint_{-\infty}^{+\infty} \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,e^{-\beta H(q,p)}\,, \\ &= \iint_{-\infty}^{+\infty} \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,e^{-\beta H_0(q,p)}\left\{1 - \frac{\beta\lambda}{8}\,q^4 + \frac{\beta^2\lambda^2}{128}\,q^8 + \cdots\right\}, \end{aligned} \tag{F.108}$$

where

$$H_0(q,p) = \frac{1}{2}\,[p^2 + \mu^2\,q^2]\,. \tag{F.109}$$

We use:

$$\int_{-\infty}^{+\infty} \mathrm{d}x\,e^{-\beta x^2/2}x^{2n} = \frac{1\cdot 3\cdot 5\cdots(2n-1)}{\beta^n}\sqrt{\frac{2\pi}{\beta}}\,. \tag{F.110}$$

So

$$Z_0(\beta) = \iint_{-\infty}^{+\infty} \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi\hbar}\,e^{-\beta H_0(q,p)} = \frac{1}{\beta\hbar\mu} = \frac{\mu^3}{\lambda\hbar}\left(\frac{\lambda}{\beta\mu^4}\right)\,. \tag{F.111}$$

and (F.108) becomes:

$$Z(\beta) = Z_0(\beta)\left\{1 - \frac{3}{2^3}\,\frac{\lambda}{\beta\mu^4} + \frac{3\cdot 5\cdot 7}{2\cdot 2^7}\left(\frac{\lambda}{\beta\mu^4}\right)^2 + \cdots\right\}. \tag{F.112}$$

This is not a high temperature expansion!

In terms of the universal function $I(\beta')$, we have:

$$I(\beta) = \iint_{-\infty}^{+\infty} \frac{\mathrm{d}q\,\mathrm{d}p}{2\pi}\,\exp\left\{-\beta\left\{\frac{1}{2}[p^2+q^2]+\frac{1}{8}\,q^4\right\}\right\} = \frac{1}{\sqrt{2\pi\beta}}\int_{-\infty}^{+\infty}\mathrm{d}q\,\exp\left\{-\beta\left\{\frac{1}{2}q^2+\frac{1}{8}\,q^4\right\}\right\}. \quad \text{(F.113)}$$

Clearly $I(\beta) \to +\infty$ as $\beta \to 0$. We can numerically integrate (F.113) to find the exact function.

References

[1] P. C. Martin, E. D. Siggia, and H. A. Rose, Phys. Rev. A **8**, 423 (1973).

[2] H. A. Rose, Ph.D. thesis, Harvard University, Cambridge, MA (1974).

Appendix G

Schwinger's Boson calculus theory of angular momentum

In a paper in 1952, Schwinger [1] invented a harmonic oscillator basis for angular momentum eigenvectors. The method consists of introducing two Bosonic harmonic oscillators and a connection between the Bosonic operators and the angular momentum algebra. Now we recognize this procedure as a second quantization of the quantum mechanical angular momentum operator, using spinors, but with Bose statistics. Of course, no *real* oscillators are involved, only the creation and annihilation features of the Boson operators are used, as in second quantization. We call the eigenvectors of the Bosonic oscillators a *Bosonic basis* for the angular momentum eigenvectors.

Schwinger's method is an explicit formulation of the group theory approach by Wigner[2][p. 163], which employed the isomorphism between the $SO(3)$ and $SU(2)$ groups, and can be regarded as a useful formalism to compute the rotation matrices and Clebsch-Gordan coefficients. In this section, we explain Schwinger's remarkable theory of angular momentum and derive some formulas for the rotation matrices and Clebsch-Gordan coefficients using his methods.

G.1 Boson calculus

Let us define two sets of independent creation and annihilation operators $A_\pm^\dagger$ and $A_\pm$, which satisfy the commutation relations:

$$[\,A_m, A_{m'}^\dagger\,] = \delta_{m,m'}\,, \qquad [\,A_m, A_{m'}\,] = 0\,, \qquad [\,A_m^\dagger, A_{m'}^\dagger\,] = 0\,, \tag{G.1}$$

for $m = \pm 1/2$. Common eigenvectors of $N_\pm = A_\pm^\dagger\, A_\pm$ are given by $|\,n_+, n_-\,)$ which satisfy the eigenvalue equation:[1]

$$N_+\,|\,n_+, n_-\,) = n_+\,|\,n_+, n_-\,)\,, \qquad N_-\,|\,n_+, n_-\,) = n_-\,|\,n_+, n_-\,)\,, \tag{G.2}$$

with the occupation number eigenvalues given by the non-negative integers: $n_\pm = 0, 1, 2, \ldots$. Also

$$\begin{aligned} A_+\,|\,n_+, n_-\,) &= \sqrt{n_+}\,|\,n_+ - 1, n_-\,\rangle\,, & A_-\,|\,n_+, n_-\,) &= \sqrt{n_-}\,|\,n_+, n_- - 1\,)\,, \\ A_+^\dagger\,|\,n_+, n_-\,) &= \sqrt{n_+ + 1}\,|\,n_+ + 1, n_-\,)\,, & A_-^\dagger\,|\,n_+, n_-\,) &= \sqrt{n_- + 1}\,|\,n_+, n_- + 1\,)\,. \end{aligned}$$

Next we define a two-dimensional column matrix A by:

$$A = \begin{pmatrix} A_+ \\ A_- \end{pmatrix}\,, \qquad A^\dagger = \begin{pmatrix} A_+^\dagger\,, & A_-^\dagger \end{pmatrix}\,, \tag{G.3}$$

[1] In order to avoid confusion between basis sets, we designate occupation number vectors by $|\,n_+, n_-\,)$.

and construct an angular momentum vector **J** using the classical two-dimentional Pauli matrices:

$$\mathbf{J} = \frac{\hbar}{2}\, A^\dagger\, \boldsymbol{\sigma}\, A = \frac{\hbar}{2}\, \left(A_+^\dagger\,,\quad A_-^\dagger\right) \begin{pmatrix} +\hat{\mathbf{e}}_z & \hat{\mathbf{e}}_x - i\hat{\mathbf{e}}_y \\ \hat{\mathbf{e}}_x + i\hat{\mathbf{e}}_y & -\hat{\mathbf{e}}_z \end{pmatrix} \begin{pmatrix} A_+ \\ A_- \end{pmatrix}. \tag{G.4}$$

Explicitly, we find:

$$J_x = \frac{\hbar}{2}\,\left(A_+^\dagger\, A_- + A_-^\dagger\, A_+\right), \qquad J_y = \frac{\hbar}{2i}\left(A_+^\dagger\, A_- - A_-^\dagger\, A_+\right), \qquad J_z = \frac{\hbar}{2}\left(A_+^\dagger\, A_+ - A_-^\dagger\, A_-\right),$$
$$J_+ = J_x + iJ_y = \hbar\, A_+^\dagger\, A_-\,, \qquad J_- = J_x - iJ_y = \hbar\, A_-^\dagger\, A_+\,.$$

It is easy to show that these definitions obey the commutation rules for angular momentum:

$$[\, J_i, J_j\,] = i\hbar\,\epsilon_{ijk} J_k\,. \tag{G.5}$$

Biedenharn [3][p. 214] called Eq. (G.4) the **Jordan-Schwinger** map, which is a mapping of Bosonic operators to the angular momentum operator **J**, linear with respect to $A_\pm$ and anti-linear with respect to $A_\pm^\dagger$.

Exercise 87. Prove Eqs. (G.5).

Operating on an eigenvector $|\, n_+, n_-\,)$ by J_z gives:

$$J_z\,|\, n_+, n_-\,) = \frac{\hbar}{2}\left(A_+^\dagger\, A_+ - A_-^\dagger\, A_-\right)|\, n_+, n_-\,) = \frac{\hbar}{2}\left(n_+ - n_-\right)|\, n_+, n_-\,) \equiv \hbar\, m\,|\, n_+, n_-\,)\,, \tag{G.6}$$

so $m = (\, n_+ - n_-\,)/2$. From Eq. (21.4), J^2 can be written as:

$$\begin{aligned} J^2 &= \frac{1}{2}\,(J_- J_+ + J_+ J_-) + J_z^2 \\ &= \left(\frac{\hbar}{2}\right)^2 \left\{ 2\left(A_-^\dagger\, A_+\, A_+^\dagger\, A_- + A_+^\dagger\, A_-\, A_-^\dagger\, A_+\right) + \left(A_+^\dagger\, A_+ - A_-^\dagger\, A_-\right)^2 \right\} \end{aligned} \tag{G.7}$$

So operating on an eigenvector $|\, n_+, n_-\,\rangle$ by J^2 gives:

$$\begin{aligned} J^2\,|\, n_+, n_-\,) &= \left(\frac{\hbar}{2}\right)^2 \left\{ 2\left(A_-^\dagger\, A_+\, A_+^\dagger\, A_- + A_+^\dagger\, A_-\, A_-^\dagger\, A_+\right) + \left(A_+^\dagger\, A_+ - A_-^\dagger\, A_-\right)^2 \right\} |\, n_+, n_-\,) \\ &= \left(\frac{\hbar}{2}\right)^2 \left\{ 2\left(n_-\,(n_+ + 1) + n_+\,(n_- + 1)\right) + \left(n_+ - n_-\right)^2 \right\} |\, n_+, n_-\,) \\ &= \hbar^2 \left\{ \left(\frac{n_+ + n_-}{2}\right)^2 + \left(\frac{n_+ + n_-}{2}\right) \right\} |\, n_+, n_-\,) \equiv \hbar^2\, j(\, j+1\,)\,|\, n_+, n_-\,)\,. \end{aligned} \tag{G.8}$$

So $j = (n_+ + n_-)/2$, and we find that $n_\pm = j \pm m$. We also see that:

$$\begin{aligned} J_+\,|\, n_+, n_-\,) &= \hbar\,\sqrt{(n_+ + 1)\, n_-}\,|\, n_+ + 1, n_- - 1\,)\,, \\ J_-\,|\, n_+, n_-\,) &= \hbar\,\sqrt{n_+\,(n_- + 1)}\,|\, n_+ - 1, n_- + 1\,)\,. \end{aligned} \tag{G.9}$$

So, when acting on the vectors $|\, j+m, j-m\,)$, gives:

$$\begin{aligned} J_\pm\,|\, j, m\,) = J_\pm\,|\, j+m, j-m\,) &= \hbar\,\sqrt{(j \mp m)\,(j \pm m + 1)}\,|\, j+m\pm 1, j-m\mp 1\,) \\ &= \hbar\, A(j, \mp m)\,|\, j, m \pm 1\,\rangle\,, \end{aligned} \tag{G.10}$$

in agreement with Eqs. (21.2) and (21.3). We also note that

$$J_+\,|\, 2j, 0\,) = 0\,, \qquad \text{and} \qquad J_-\,|\, 0, 2j\,) = 0\,. \tag{G.11}$$

Normalized angular momentum eigenvectors are then given in terms of the occupation number basis by:

$$|\,j,m\,) \equiv |\,j+m,j-m\,) = \frac{(\,A_+^\dagger\,)^{j+m}(\,A_-^\dagger\,)^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}\,|\,0\,)\,. \tag{G.12}$$

Eq. (G.12) gives angular momentum eigenvectors for any value of j and m in terms of two creation operators acting on a ground state $|\,0\,)$.[2]

Exercise 88. Find the occupation number vectors $|\,n_+,n_-\,)$ for $j=1/2$ and $j=1$.

G.2 Connection to quantum field theory

In the section, we show the connection between second quantized field operators and the Boson calculus. Let us define a two-component field operator $\Psi_m(x)$ with $m=\pm$ as:

$$\Psi_m(\mathbf{r}) = \sum_{k,q} A_{k,q}\,[\,\psi_{k,q}(\mathbf{r})\,]_m\,, \tag{G.13}$$

where $\psi_{k,q}(\mathbf{r})$ is a two-component wave function and $A_{k,q}$ an operator. The wave functions $\psi_{k,q}(\mathbf{r})$ satisfy the orthogonal and completeness relations:

$$\begin{aligned}\sum_m \int \mathrm{d}^3x\,[\,\psi^*_{k,q}(\mathbf{r})\,]_m\,[\,\psi_{k',q'}(\mathbf{r})\,]_m &= \delta_{k,k'}\,\delta_{m,m'}\,,\\ \sum_{k,q}[\,\psi_{k,q}(\mathbf{r})\,]_m\,[\,\psi^*_{k,q}(\mathbf{r}')\,]_{m'} &= \delta_{m,m'}\,\delta(\mathbf{r}-\mathbf{r}')\,.\end{aligned} \tag{G.14}$$

If we choose the $A_{k,q}$ operators to satisfy the commutation relation:

$$[\,A_{k,q},A^\dagger_{k',q'}\,] = \delta_{k,k'}\,\delta_{q,q'}\,, \tag{G.15}$$

then the field operators $\Psi_m(\mathbf{r})$ satisfy Bose statistics:

$$[\,\Psi_m(\mathbf{r}),\Psi^\dagger_{m'}(\mathbf{r}')\,] = \delta_{m,m'}\,\delta(\mathbf{r}-\mathbf{r}')\,. \tag{G.16}$$

On the other hand, if we choose the $A_{k,q}$ operators to satisfy the anticommutation relation:

$$\{\,A_{k,q},A^\dagger_{k',q'}\,\} = \delta_{k,k'}\,\delta_{q,q'}\,, \tag{G.17}$$

then the field operators $\Psi_m(\mathbf{r})$ satisfy Fermi statistics:

$$\{\,\Psi_m(\mathbf{r}),\Psi^\dagger_{m'}(\mathbf{r}')\,\} = \delta_{m,m'}\,\delta(\mathbf{r}-\mathbf{r}')\,. \tag{G.18}$$

Now let us take basis wave functions of the form: $\psi_{k,q}(\mathbf{r}) = \psi_k(\mathbf{r})\,\chi_q$, where χ_q are the two-component eigenspinors of σ_z. Then the field operators become:

$$\Psi_m(\mathbf{r}) = \sum_{k,q} A_{k,q}\,\psi_k(\mathbf{r})\,[\,\chi_q\,]_m = \sum_{k,q}[\,A_k\,]_m\,\psi_k(\mathbf{r})\,, \tag{G.19}$$

where A_k is the two-component operator:

$$A_k = \sum_q A_{k,q}\,\chi_q = \begin{pmatrix} A_{k,+}\\ A_{k,-}\end{pmatrix}\,. \tag{G.20}$$

[2] We use $|\,0\,)$ to designate the ground state with $n_+=n_-=0$.

Now we can define the angular momentum in this field by:

$$\mathbf{J} = \frac{\hbar}{2}\int \mathrm{d}^3x\,\Psi^\dagger(\mathbf{r})\,\boldsymbol{\sigma}\,\Psi\ (\mathbf{r}) = \frac{\hbar}{2}\sum_k A_k^\dagger\,\boldsymbol{\sigma}\,A_k\,. \tag{G.21}$$

For the case of only one state $k=1$, this is the same Jordan-Schwinger map of Eq. (G.4). However, here we see that we can choose the operators A_k to obey either commutators and Bose statistics or anticommutators and Fermi statistics. Schwinger chose these to be commutators, but we could equally take them to be anticommutators. We leave it to an exercise to show that if we choose Fermi statistics, the angular momentum operator $\mathbf{J}$ still obeys the correct angular momentum commutation relations.

For the Bose case, we can think of the vectors described by Eq. (G.12) as being made up of primitive spin-1/2 "particles" with $n_+ = j+m$ spin up in the z-direction and $n_- = j-m$ spin down in the z-direction. That is, we can think of the system as composed of $2j$ spin-1/2 *Bose* particles.

Exercise 89. Show that if A_m obeys anticommutation relations (Fermi statistics), the second quantized angular momentum vector $\mathbf{J}$, defined by Eq. (G.21) obeys the usual commutation relation: $[\,J_i, J_j\,] = i\hbar\,\epsilon_{ijk}\,J_k$.

G.3 Hyperbolic vectors

The eigenvectors $J_\pm$ defined above are ladder operators which, when acting on the states $|\,j,m\,\rangle$ create states of $|\,j,m\pm1\,\rangle$ for *fixed* values of j. We can also find operators which, when operating on $|\,j,m\,\rangle$ create states of $|\,j\pm1,m\,\rangle$ for *fixed* values of m. That is, the role of j and m is reversed. To this end, we define the vector operator $\mathbf{K}$ by:

$$K_x = \frac{\hbar}{2}\left(A_+^\dagger A_-^\dagger + A_+\,A_-\right), \quad K_y = \frac{\hbar}{2i}\left(A_+^\dagger A_-^\dagger - A_+\,A_-\right), \quad K_z = \frac{\hbar}{2}\left(A_+^\dagger A_+ + A_-^\dagger A_- + 1\right),$$
$$K_+ = K_x + iK_y = \hbar\,A_+^\dagger A_-^\dagger\,, \qquad K_- = K_x - iK_y = \hbar\,A_+\,A_-\,. \tag{G.22}$$

With these definitions, it is easy to show that the K_i operators obey the commutation rules:

$$[\,K_i, K_j\,] = -i\hbar\,\epsilon_{ijk}K_k\,, \tag{G.23}$$

which are the commutation rules for an angular momentum with the sign reversed. The "length" of the vector $\mathbf{K}$ in this hyperbolic space is defined by:

$$K^2 = K_x^2 + K_y^2 - K_z^2 = \frac{1}{2}\left(K_+\,K_- + K_-\,K_-\right) - K_z^2\,, \tag{G.24}$$

and is *not* positive definite. Schwinger [1] calls these operators hyperbolic because of the reversed sign. One can easily check that

$$[\,J_z, K_z\,] = 0\,, \tag{G.25}$$

and so J_z and K_z have common eigenvectors. The eigenvector of K_z is given by:

$$K_z\,|\,j,m\,\rangle = K_z\,|\,j+m, j-m\,) = \frac{\hbar}{2}\left(j+m+j-m+1\right)|\,j+m, j-m\,) = \hbar\left(j+1/2\right)|\,j,m\,\rangle\,. \tag{G.26}$$

For $K_\pm$ we find:

$$\begin{aligned} K_+\,|\,j,m\,\rangle &= \hbar\,A_+^\dagger A_-^\dagger\,|\,j+m, j-m\,) = \hbar\sqrt{(j+m+1)\,(j-m+1)}\,|\,j+m+1, j-m+1\,) \\ &= \hbar\sqrt{(j+m+1)\,(j-m+1)}\,|\,j+1,m\,\rangle\,, \\ K_-\,|\,j,m\,\rangle &= \hbar\,A_+\,A_-\,|\,j+m, j-m\,) = \hbar\sqrt{(j+m)\,(j-m)}\,|\,j+m-1, j-m-1\,) \\ &= \hbar\sqrt{(j+m)\,(j-m)}\,|\,j-1,m\,\rangle\,, \end{aligned} \tag{G.27}$$

So $K_\pm$ create states of ± 1 additional units of j for fixed m, as we wanted. For K^2, we find:

$$\begin{aligned} K^2 &= \frac{1}{2}\left(K_+ K_- + K_- K_+\right) - K_z^2 \\ &= \left(\frac{\hbar}{2}\right)^2 \left\{ 2\left(A_+^\dagger A_-^\dagger A_+ A_- + A_+ A_- A_+^\dagger A_-^\dagger\right) - \left(A_+^\dagger A_+ + A_-^\dagger A_- + 1\right)^2 \right\}. \end{aligned} \tag{G.28}$$

Operating on an eigenvector $|\, n_+, n_-\,)$ by K^2 gives:

$$\begin{aligned} K^2\,|\, n_+, n_-\,) &= \left(\frac{\hbar}{2}\right)^2 \left\{ 2\left(A_+^\dagger A_-^\dagger A_+ A_- + A_+ A_- A_+^\dagger A_-^\dagger\right) - \left(A_+^\dagger A_+ + A_-^\dagger A_- + 1\right)^2 \right\} |\, n_+, n_-\,) \\ &= \left(\frac{\hbar}{2}\right)^2 \left\{ 2\left(n_+ n_- + (n_+ + 1)(n_- + 1)\right) - \left(n_+ + n_- + 1\right)^2 \right\} |\, n_+, n_-\,) \\ &= \hbar^2 \left\{ \frac{1}{4} - \left(\frac{n_+ - n_-}{2}\right)^2 \right\} |\, n_+, n_-\,) = \left\{ \frac{\hbar^2}{4} - J_z^2 \right\} |\, n_+, n_-\,). \end{aligned} \tag{G.29}$$

That is

$$K^2\,|\, j, m\,\rangle = \hbar^2 \left\{ \frac{1}{4} - m^2 \right\} |\, j, m\,\rangle. \tag{G.30}$$

So K^2, as we have defined it, is diagonal in the $|\, j, m\,\rangle$ basis, and is simply related to J_z.

G.4 Coherent states

It will be useful to study coherent states of these occupation number basis vectors of angular momentum. Following the development of coherent states in Section 16.4, we define these states by

$$A_m\,|\, a_+, a_-\,) = a_m\,|\, a_+, a_-\,), \tag{G.31}$$

where $a_\pm$ are complex numbers. So from Eq. (G.12), we find:

$$(\, a_+, a_-\,|\, j, m\,\rangle = (\, a_+, a_-\,|\, j+m, j-m\,) = \mathcal{N}(\, a_+, a_-\,)\, \frac{\left(a_+^*\right)^{j+m} \left(a_-^*\right)^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}, \tag{G.32}$$

where $\mathcal{N}(\, a_+, a_-\,) = (\, a_+, a_-\,|\, 0, 0\,)$ is a normalization factor. The state $|\, a_+, a_-\,)$ is then given by:

$$\begin{aligned} |\, a_+, a_-\,) &= \sum_{j,m} |\, j+m, j-m\,)\,(\, j+m, j-m\,|\, a_+, a_-\,) \\ &= \mathcal{N}(\, a_+, a_-\,) \sum_{j,m} \frac{\left(a_+ A_+^\dagger\right)^{j+m} \left(a_- A_-^\dagger\right)^{j-m}}{(j+m)!\,(j-m)!}\,|\, 0\,). \end{aligned} \tag{G.33}$$

Now using the binomial theorem, we have

$$\sum_{m=-j}^{j} \frac{\left(a_+ A_+^\dagger\right)^{j+m} \left(a_- A_-^\dagger\right)^{j-m}}{(j+m)!\,(j-m)!} = \sum_{k=0}^{2j} \frac{\left(a_+ A_+^\dagger\right)^{k} \left(a_- A_-^\dagger\right)^{2j-k}}{k!\,(2j-k)!} = \left(a_+ A_+^\dagger + a_- A_-^\dagger\right)^{2j}. \tag{G.34}$$

So

$$|\, a_+, a_-\,) = \mathcal{N}(\, a_+, a_-\,)\, \exp\left\{ \sum_m a_m A_m^\dagger \right\} |\, 0\,). \tag{G.35}$$

So the coherent state is not a state of definite angular momentum, but contains all angular momentum states. In this sense, it can be thought of as a generator of all the angular momentum states. Normalizing

the coherent state, we find:

$$\begin{aligned}(\, a_+, a_- \,|\, a_+, a_- \,) &= |\,\mathcal{N}(\, a_+, a_- \,)\,|^2\,(\,0\,|\,\exp\Big\{\sum_m a_m^*\,A_m\Big\}\,\exp\Big\{\sum_{m'} a_{m'}\,A_{m'}^\dagger\Big\}\,|\,0\,) \\ &= |\,\mathcal{N}(\, a_+, a_- \,)\,|^2\,\exp\Big\{\sum_m |\,a_m\,|^2\Big\} = 1\,,\end{aligned} \tag{G.36}$$

so

$$\mathcal{N}(\, a_+, a_- \,) = \exp\Big\{-\sum_m |\,a_m\,|^2/2\Big\}\,. \tag{G.37}$$

Then

$$\begin{aligned}|\, a_+, a_- \,) &= \exp\Big\{\sum_m a_m\,A_m^\dagger - \sum_m |\,a_m\,|^2/2\Big\}\,|\,0\,) = \exp\Big\{\sum_m \big[\, a_m\,A_m^\dagger - a_m^*\,A_m\,\big]\Big\}\,|\,0\,) \\ &\equiv D(a_m, a_m^*)\,|\,0\,)\,.\end{aligned} \tag{G.38}$$

Here we have defined a unitary displacement operator $D(a_m, a_m^*)$ by:

$$D(a_m, a_m^*) = \exp\Big\{\sum_m \big[\, a_m\,A_m^\dagger - a_m^*\,A_m\,\big]\Big\}\,, \tag{G.39}$$

and has the property,

$$D^\dagger(a_m, a_m^*)\,A_m D(a_m, a_m^*) = A_m + a_m\,. \tag{G.40}$$

Example 56. A generator of J_+ can be constructed by considering the operation of $\exp\{\lambda J_+/\hbar\}$ on a coherent state. We find:

$$\begin{aligned}\exp\Big\{\lambda J_+/\hbar\Big\}\,|\, a_+, a_- \,) &= \exp\Big\{\,\lambda A_+^\dagger\,A_-\,\Big\}\,|\, a_+, a_- \,) = \exp\Big\{\,\lambda A_+^\dagger\,a_-\,\Big\}\,|\, a_+, a_- \,) \\ &= \exp\Big\{\,\lambda A_+^\dagger\,a_-\,\Big\}\,\exp\Big\{\sum_m \big[\, a_m\,A_m^\dagger - a_m^*\,A_m\,\big]\Big\}\,|\,0\,)\,.\end{aligned} \tag{G.41}$$

Again, using Eq. (B.16) and writing out these operators explicitly, we find:

$$\begin{aligned}&\exp\Big\{\lambda J_+/\hbar\Big\}\,|\, a_+, a_+^*; a_-, a_-^* \,) \\ &\qquad = \exp\Big\{\,\lambda\,a_-\,a_+^*/2\,\Big\}\,\exp\Big\{\,(\,a_+ + \lambda\,a_-\,)\,A_+^\dagger + a_-\,A_-^\dagger - a_+^*\,A_+ - a_-^*\,A_-\,\Big\}\,|\,0\,) \\ &\qquad\qquad = \exp\Big\{\,\lambda\,a_-\,a_+^*/2\,\Big\}\,|\, a_+ + \lambda\,a_-, a_+^*; a_-, a_-^* \,)\end{aligned} \tag{G.42}$$

Operating on the left by the vector $\langle\, j, m\,|$ and introducing a complete set of states gives:

$$\begin{aligned}&\sum_{m'=-j}^{j} \langle\, j, m\,|\,\exp\{\lambda J_+/\hbar\}\,|\, j, m'\,\rangle\,\langle\, j, m'\,|\, a_+, a_+^*; a_-, a_-^* \,) \\ &\qquad\qquad = e^{\,\lambda\,a_-\,a_+^*/2}\;\langle\, j, m'\,|\, a_+ + \lambda\,a_-, a_+^*; a_-, a_-^* \,)\,.\end{aligned} \tag{G.43}$$

But using the representation (G.32),

$$\langle\, j, m\,|\, a_+, a_- \,) = e^{-\sum_{m'} a_{m'}^*\,a_{m'}/2}\;\frac{(\,a_+\,)^{j+m}\,(\,a_-\,)^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}\,, \tag{G.44}$$

Eq. (G.43) becomes:

$$e^{-[\,a_+^*\,a_+ + a_-^*\,a_-\,]/2}\sum_{m'=-j}^{j}\langle\,j,m\,|\,\exp\{\lambda J_+/\hbar\}\,|\,j,m'\,\rangle\,\frac{(\,a_+\,)^{j+m'}\,(\,a_-\,)^{j-m'}}{\sqrt{(j+m')!\,(j-m')!}}$$
$$= e^{[\,\lambda\,a_-\,a_+^* - a_+^*\,(\,a_+ + \lambda\,a_-\,) - a_-^*\,a_-\,]/2}\;\frac{(\,a_+ + \lambda\,a_-\,)^{j+m}\,(\,a_-\,)^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}\,. \quad \text{(G.45)}$$

The exponential normalization factors cancel. Expanding both sides in powers of λ give:

$$\sum_{k=0}^{\infty}\sum_{m'=-j}^{j}\langle\,j,m\,|\,\big[\,J_+/\hbar\,\big]^k\,|\,j,m'\,\rangle\,\frac{\lambda^k\,(\,a_+\,)^{j+m'}\,(\,a_-\,)^{j-m'}}{k!\,\sqrt{(j+m')!\,(j-m')!}}$$
$$= \sum_{n=0}^{j+m}\frac{(j+m)!}{(j+m-n)!\,n!}\,\frac{\lambda^n\,(\,a_+\,)^{j+m-n}\,(\,a_-\,)^{j-m+n}}{\sqrt{(j+m)!\,(j-m)!}}\,. \quad \text{(G.46)}$$

Setting $n = m - m'$ on the right-hand-side of this expression gives:

$$\sum_{k=0}^{\infty}\sum_{m'=-j}^{j}\langle\,j,m\,|\,\big[\,J_+/\hbar\,\big]^k\,|\,j,m'\,\rangle\,\frac{\lambda^k\,(\,a_+\,)^{j+m'}\,(\,a_-\,)^{j-m'}}{k!\,\sqrt{(j+m')!\,(j-m')!}}$$
$$= \sum_{m'=-j}^{m}\sqrt{\frac{(j+m)!\,(j-m')!}{(j-m)!\,(j+m')!}}\,\frac{\lambda^{m-m'}\,(\,a_+\,)^{j+m'}\,(\,a_-\,)^{j-m'}}{(m-m')!\,\sqrt{(j+m')!\,(j-m')!}}\,. \quad \text{(G.47)}$$

Comparing powers of λ, we find:

$$\langle\,j,m\,|\,\big[\,J_+/\hbar\,\big]^k\,|\,j,m'\,\rangle = \delta_{k,m-m'}\sqrt{\frac{(j+m)!\,(j-m')!}{(j-m)!\,(j+m')!}} \qquad \text{for} \qquad m - m' \geq 0\,. \quad \text{(G.48)}$$

Similarly, for $[\,J_-/\hbar\,]^k$, we find:

$$\langle\,j,m\,|\,\big[\,J_-/\hbar\,\big]^k\,|\,j,m'\,\rangle = \delta_{k,m'-m}\sqrt{\frac{(j+m')!\,(j-m)!}{(j-m')!\,(j+m)!}} \qquad \text{for} \qquad m' - m \geq 0\,. \quad \text{(G.49)}$$

In particular, by setting $m' = \pm j$ in these two equations, we can easily find the vector $|\,j,m\,\rangle$ from the vectors $|\,j,\pm j\,\rangle$:

$$|\,j,m\,\rangle = \sqrt{\frac{(j-m)!}{(2j)!\,(j+m)!}}\,\big[\,J_+/\hbar\,\big]^{j+m}\,|\,j,-j\,\rangle\,,$$
$$|\,j,m\,\rangle = \sqrt{\frac{(j+m)!}{(2j)!\,(j-m)!}}\,\big[\,J_-/\hbar\,\big]^{j-m}\,|\,j,+j\,\rangle\,. \quad \text{(G.50)}$$

G.5 Rotation matrices

In this section, we use Boson second quantized basis states to find useful formulas for the $D^{(j)}_{m,m'}(R)$ rotation matrices. It is closely related to the use of Cayley-Klein parameters which we discussed in Section 21.2.4.

In the Boson calculus, the unitary operator for rotations is given by:

$$U(R) = \exp[\,i\,\hat{\mathbf{n}}\cdot\mathbf{J}\,\theta/\hbar\,] = \exp[\,i\,A^\dagger\,\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\,A\,\theta/2\,]\,, \quad \text{(G.51)}$$

where we have used Eq. (G.4). Note that $U(R)$ is an *operator* in the second quantized basis, and that $U(R)\,|\,0\,\rangle = |\,0\,\rangle$. One can easily show that the spinor A transforms as (see Exercise 90 below):

$$U^\dagger(R)\,A_m\,U(R) = \exp[\,-i\,A^\dagger\,\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\,A\,\theta/2\,]\,A_m\,\exp[\,i\,A^\dagger\,\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\,A\,\theta/2\,] = D_{m,m'}(R)\,A_{m'}\,, \tag{G.52}$$

where the 2×2 rotation matrix $D_{m,m'}(R)$ can be labeled by any of the parameterizations given in Theorem **??**: Cayley-Klein, axis and angle of rotation, or Euler angles. That is:

$$\begin{aligned} D(R) &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \cos(\theta/2) + i\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,\sin(\theta/2) \\ &= \begin{pmatrix} e^{i(+\gamma+\alpha)/2}\,\cos(\beta/2) & e^{i(+\gamma-\alpha)/2}\,\sin(\beta/2) \\ -e^{i(-\gamma+\alpha)/2}\,\sin(\beta/2) & e^{i(-\gamma-\alpha)/2}\,\cos(\beta/2) \end{pmatrix}. \end{aligned} \tag{G.53}$$

This D-matrix is identical to the $D^{(1/2)}_{m,m'}(R)$ matrix for spin 1/2. Operators which transformation according to Eq. (G.52) are called rank one-half tensors.

Exercise 90. Using Eq. (B.14) in Appendix B, compute out a few terms in the expansion to convince yourself of the truth of Eq. (G.52).

Using Eq. (**??**), we find:

$$\begin{aligned} U^\dagger(R)\,J_i\,U(R) &= U^\dagger(R)\,\frac{\hbar}{2}\,A^\dagger\,\sigma_i\,A\,U(R) = \frac{\hbar}{2}\,U^\dagger(R)\,A^\dagger\,U(R)\,\sigma_i\,U^\dagger(R)\,A\,U(R) \\ &= \frac{\hbar}{2}\,A^\dagger\,D^\dagger(R)\,\sigma_i\,D(R)\,A = R_{ij}\,\frac{\hbar}{2}\,A^\dagger\,\sigma_j\,A = R_{ij}\,J_j \equiv J'_i\,, \end{aligned} \tag{G.54}$$

as required. We now prove the following theorem:

Theorem 72 (D-matrix)**.** *The $D^{(j)}_{m,m'}(R)$ matrix for any j is given by:*

$$\begin{aligned} D^{(j)}_{m,m'}(R) &= \sqrt{(j+m)!\,(j-m)!\,(j+m')\,(j-m')} \\ &\times \sum_{s=0}^{j+m}\sum_{r=0}^{j-m}\delta_{s-r,m-m'}\,\frac{\big(\,D_{+,+}(R)\,\big)^{j+m-s}\,\big(\,D_{+,-}(R)\,\big)^{s}\,\big(\,D_{-,+}(R)\,\big)^{r}\,\big(\,D_{-,-}(R)\,\big)^{j-m-r}}{s!\,(j+m-s)!\,r!\,(j-m-r)!}\,, \end{aligned} \tag{G.55}$$

where the elements of the two-dimensional matrix $D(R)$ are given by Eq. (G.53) with the rows and columns labeled by $\pm$. Eq. (G.55) relates the D-matrices for any j to the D-matrices for $j = 1/2$. The range of the sums over s and r is determined by the values of j, m, and m'.

Proof. The occupation number states, transform according to:

$$\begin{aligned} |\,j,m\,\rangle' = |\,j+m,j-m\,)' &= U^\dagger(R)\,|\,j+m,j-m\,) = U^\dagger(R)\,\frac{\big(\,A^\dagger_+\,\big)^{j+m}\big(\,A^\dagger_-\,\big)^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}\,|\,0\,) \\ &= \frac{\big(\,U^\dagger(R)\,A^\dagger_+\,U(R)\,\big)^{j+m}\,\big(\,U^\dagger(R)\,A^\dagger_-\,U(R)\,\big)^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}\,|\,0\,) \\ &= \frac{\big(\,D^*_{+,+}(R)\,A^\dagger_+ + D^*_{+,-}(R)\,A^\dagger_-\,\big)^{j+m}\,\big(\,D^*_{-,+}(R)\,A^\dagger_+ + D^*_{-,-}(R)\,A^\dagger_-\,\big)^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}\,|\,0\,) \end{aligned} \tag{G.56}$$

Using the binomial expansion, we have:

$$\big(\,D^*_{+,+}(R)\,A^\dagger_+ + D^*_{+,-}(R)\,A^\dagger_-\,\big)^{j+m} = \sum_{s=0}^{j+m}\frac{(j+m)!}{s!\,(j+m-s)!}\,\big(\,D^*_{+,+}(R)\,A^\dagger_+\,\big)^{j+m-s}\,\big(\,D^*_{+,-}(R)\,A^\dagger_-\,\big)^{s}\,,$$

$$\big(\,D^*_{-,+}(R)\,A^\dagger_+ + D^*_{-,-}(R)\,A^\dagger_-\,\big)^{j-m} = \sum_{r=0}^{j-m}\frac{(j-m)!}{r!\,(j-m-r)!}\,\big(\,D^*_{-,+}(R)\,A^\dagger_+\,\big)^{r}\,\big(\,D^*_{-,-}(R)\,A^\dagger_-\,\big)^{j-m-r}\,.$$

So Eq. (G.56) becomes:

$$
\begin{aligned}
|\,j,m\,\rangle' &= \sqrt{(j+m)!\,(j-m)!} \\
&\times \sum_{s=0}^{j+m}\sum_{r=0}^{j-m} \frac{\big(\,D^*_{+,+}(R)\,\big)^{j+m-s}\,\big(\,D^*_{+,-}(R)\,\big)^{s}\,\big(\,D^*_{-,+}(R)\,\big)^{r}\,\big(\,D^*_{-,-}(R)\,\big)^{j-m-r}}{s!\,(j+m-s)!\,r!\,(j-m-r)!}\,\big(\,A^\dagger_+\,\big)^{j+m-s+r}\,\big(\,A^\dagger_-\,\big)^{j-m+s-r}\,|\,0\,\rangle \\
&= \sum_{m'=-j}^{j}\Bigg\{\sum_{s=0}^{j+m}\sum_{r=0}^{j-m}\delta_{s-r,m-m'}\,\sqrt{(j+m)!\,(j-m)!\,(j+m')\,(j-m')} \\
&\times \frac{\big(\,D^*_{+,+}(R)\,\big)^{j+m-s}\,\big(\,D^*_{+,-}(R)\,\big)^{s}\,\big(\,D^*_{-,+}(R)\,\big)^{r}\,\big(\,D^*_{-,-}(R)\,\big)^{j-m-r}}{s!\,(j+m-s)!\,r!\,(j-m-r)!}\Bigg\}\,\frac{\big(\,A^\dagger_+\,\big)^{j+m'}\big(\,A^\dagger_-\,\big)^{j-m'}}{\sqrt{(j+m')!\,(j-m')!}}\,|\,0\,\rangle \\
&= \sum_{m'=-j}^{j} D^{(j)*}_{m,m'}(R)\,|\,j,m'\,\rangle\,, \qquad \text{(G.57)}
\end{aligned}
$$

where

$$
\begin{aligned}
D^{(j)}_{m,m'}(R) &= \sqrt{(j+m)!\,(j-m)!\,(j+m')\,(j-m')} \\
&\times \sum_{s=0}^{j+m}\sum_{r=0}^{j-m}\delta_{s-r,m-m'}\,\frac{\big(\,D_{+,+}(R)\,\big)^{j+m-s}\,\big(\,D_{+,-}(R)\,\big)^{s}\,\big(\,D_{-,+}(R)\,\big)^{r}\,\big(\,D_{-,-}(R)\,\big)^{j-m-r}}{s!\,(j+m-s)!\,r!\,(j-m-r)!}\,, \qquad \text{(G.58)}
\end{aligned}
$$

which is what we were trying to prove. □

It is easy to check that for $j = 1/2$ we get the correct result. It is useful to work out a special case for the Euler angle description of the rotation when $\alpha = \gamma = 0$. Then

$$
D_{m,m'}(0,\beta,0) = \begin{pmatrix} \cos(\beta/2) & \sin(\beta/2) \\ -\sin(\beta/2) & \cos(\beta/2) \end{pmatrix}\,, \qquad \text{(G.59)}
$$

and (G.55) becomes:

$$
\begin{aligned}
D^{(j)}_{m,m'}(0,\beta,0) &= \sqrt{(j+m)!\,(j-m)!\,(j+m')\,(j-m')} \\
&\times \sum_{s=0}^{j+m}\sum_{r=0}^{j-m}\delta_{s-r,m-m'}\,\frac{(-)^r\,\big(\,\cos(\beta/2)\,\big)^{2j-s-r}\,\big(\,\sin(\beta/2)\,\big)^{s+r}}{s!\,(j+m-s)!\,r!\,(j-m-r)!} \\
&= \sqrt{(j+m)!\,(j-m)!\,(j+m')\,(j-m')} \\
&\times \sum_{\sigma}\frac{(-)^{j-\sigma-m}\,\big(\,\cos(\beta/2)\,\big)^{2\sigma+m+m'}\,\big(\,\sin(\beta/2)\,\big)^{2j-2\sigma-m-m'}}{\sigma!\,(j-\sigma-m)!\,(j-\sigma-m')!\,(\sigma+m+m')!}\,,
\end{aligned} \qquad \text{(G.60)}
$$

in agreement with Edmonds [4][Eq. (4.1.15), p. 57]. In the last line we put $s = j - m - \sigma$, so σ is an integer.

We also can work out some properties of the D-functions using this method. First, we note that for $\alpha = \gamma = 0$, $\beta = \pi$,

$$
D_{m,m'}(0,\pi,0) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\,. \qquad \text{(G.61)}
$$

So in this case, Eq. (G.56) becomes:

$$
\begin{aligned}
|\,j,m\,\rangle' &= \frac{\big(\,A^\dagger_-\,\big)^{j+m}\,\big(\,-A^\dagger_+\,\big)^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}\,|\,0\,\rangle = (-)^{j-m}\,\frac{\big(\,A^\dagger_+\,\big)^{j-m}\big(\,A^\dagger_-\,\big)^{j+m}}{\sqrt{(j+m)!\,(j-m)!}}\,|\,0\,\rangle \\
&= \sum_{m'=-j}^{j}(-)^{j-m}\,\delta_{m,-m'}\,\frac{\big(\,A^\dagger_+\,\big)^{j+m'}\big(\,A^\dagger_-\,\big)^{j-m'}}{\sqrt{(j+m')!\,(j-m')!}}\,|\,0\,\rangle = \sum_{m'=-j}^{j}(-)^{j-m}\,\delta_{m,-m'}\,|\,j,m'\,\rangle\,,
\end{aligned} \qquad \text{(G.62)}
$$

So

$$D^{(j)}_{m,m'}(0,\pi,0) = (-)^{j-m}\,\delta_{m,-m'}\,, \qquad \text{similarly,} \qquad D^{(j)}_{m,m'}(0,-\pi,0) = (-)^{j+m}\,\delta_{m,-m'}\,. \tag{G.63}$$

The D-matrices can also be calculated directly in the occupation number basis. From Eq. (G.12), we have:

$$\begin{aligned} D^{(j)}_{m,m'}(R) &= \langle\, j,m\,|\,U(R)\,|\,j,m'\,\rangle = (\,j+m,j-m\,|\,U(R)\,|\,j+m',j-m'\,) \\ &= (\,0\,|\,\frac{(\,A_+\,)^{j+m}(\,A_-\,)^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}\,U(R)\,\frac{(\,A^\dagger_+\,)^{j+m'}(\,A^\dagger_-\,)^{j-m'}}{\sqrt{(j+m')!\,(j-m')!}}\,|\,0\,) \\ &= \frac{(\,0\,|\,U^\dagger(R)\,(\,A_+\,)^{j+m}(\,A_-\,)^{j-m}\,U(R)\,(\,A^\dagger_+\,)^{j+m'}(\,A^\dagger_-\,)^{j-m'}\,|\,0\,)}{\sqrt{(j+m)!\,(j-m)!\,(j+m')!\,(j-m')!}} \\ &= \frac{(\,0\,|\,(\,A'_+(R)\,)^{j+m}(\,A'_-(R)\,)^{j-m}\,(\,A^\dagger_+\,)^{j+m'}(\,A^\dagger_-\,)^{j-m'}\,|\,0\,)}{\sqrt{(j+m)!\,(j-m)!\,(j+m')!\,(j-m')!}}\,, \end{aligned} \tag{G.64}$$

where

$$\begin{aligned} A'_+(R) &= D_{+,+}(R)\,A_+ + D_{+,-}(R)\,A_-\,, \\ A'_-(R) &= D_{-,+}(R)\,A_+ + D_{-,-}(R)\,A_-\,. \end{aligned} \tag{G.65}$$

The method to be used here is to move the creation operators to the left and the annihilation operators to the right, using the commutation properties, so that they operator on the ground state $|\,0\,)$ and give zero.

Exercise 91. Use Eq. (G.64) to find the components of $D^{(j)}_{m,m'}(0,\beta,0)$ for $j = 1/2$ and $j = 1$.

G.6 Addition of angular momentum

In this section, we show how to use Bose operators to construct an eigenvector of total angular momentum which is the sum of two angular momentum systems. We will use this result to find Clebsch-Gordan coefficients, and a generating function for these coefficients.

So let $A^\dagger_{1,m}$ and $A^\dagger_{2,m}$ be two commuting sets of creation operators, $m = \pm$, obeying the algebra:

$$[\,A_{\alpha,m},A^\dagger_{\beta,m'}\,] = \delta_{m,m'}\delta_{\alpha,\beta}\,, \qquad [\,A_{\alpha,m},A_{\beta,m'}\,] = [\,A^\dagger_{\alpha,m},A^\dagger_{\beta,m'}\,] = 0\,, \tag{G.66}$$

with $\alpha,\beta = (1,2)$, and describing the two angular momentum system by the Jordan-Schwinger maps:

$$\mathbf{J}_1 = \frac{\hbar}{2}\,A^\dagger_1\,\boldsymbol{\sigma}\,A_1\,, \qquad \mathbf{J}_2 = \frac{\hbar}{2}\,A^\dagger_2\,\boldsymbol{\sigma}\,A_2\,. \tag{G.67}$$

Eigenvectors of the number operators $N_{\alpha,m} = A^\dagger_{\alpha,m}\,A_{\alpha,m}$ are written in a shorthand notation as $|\,n_{\alpha,m}\,\rangle \equiv |\,n_{1,+},n_{1,-},n_{2,+},n_{2,-}\,\rangle$ and satisfy:

$$\begin{aligned} N_{\alpha,m}\,|\,n_{\alpha,m}\,) &= n_{\alpha,m}\,|\,n_{\alpha,m}\,)\,, \\ A_{\alpha,m}\,|\,n_{\alpha,m}\,) &= \sqrt{n_{\alpha,m}}\,|\,n_{\alpha,m}-1\,) \\ A^\dagger_{\alpha,m}\,|\,n_{\alpha,m}\,) &= \sqrt{n_{\alpha,m}+1}\,|\,n_{\alpha,m}+1\,)\,, \end{aligned} \tag{G.68}$$

with $n_{\alpha,m} = 0,1,2,\ldots$. We put the $A^\dagger_{\alpha,m}$ operators into a 2×2 matrix of the form:

$$A^\dagger = \begin{pmatrix} A^\dagger_{1,+} & A^\dagger_{1,-} \\ A^\dagger_{2,+} & A^\dagger_{2,-} \end{pmatrix}\,, \qquad \text{so that} \qquad A = \begin{pmatrix} A_{1,+} & A_{2,+} \\ A_{1,-} & A_{2,-} \end{pmatrix}\,. \tag{G.69}$$

Note that $A^\dagger$ has rows and columns labeled by (α, m) but that A has rows and columns labeled by (m, α). The total angular momentum $\mathbf{J}$ is given by the mapping:

$$\mathbf{J} = \frac{\hbar}{2}\,\mathrm{Tr}[\,A^\dagger\,\boldsymbol{\sigma}\,A\,] = \frac{\hbar}{2}\sum_{m,m',\alpha} A^\dagger_{\alpha,m}\,\boldsymbol{\sigma}_{m,m'}\,A_{\alpha,m'} = \mathbf{J}_1 + \mathbf{J}_2\,. \tag{G.70}$$

Explicitly, we find:

$$\begin{aligned} J_x &= \frac{\hbar}{2}\sum_\alpha\big(\,A^\dagger_{\alpha,+}\,A_{\alpha,-} + A^\dagger_{\alpha,-}\,A_{\alpha,+}\,\big)\,,\\ J_y &= \frac{\hbar}{2i}\sum_\alpha\big(\,A^\dagger_{\alpha,+}\,A_{\alpha,-} - A^\dagger_{\alpha,-}\,A_{\alpha,+}\,\big)\,,\\ J_z &= \frac{\hbar}{2}\sum_\alpha\big(\,A^\dagger_{\alpha,+}\,A_{\alpha,+} - A^\dagger_{\alpha,-}\,A_{\alpha,-}\,\big) = \frac{\hbar}{2}\,\big(\,N_+ - N_-\,\big)\,, \end{aligned} \tag{G.71}$$

and

$$J_+ = J_x + iJ_y = \hbar\sum_\alpha A^\dagger_{\alpha,+}\,A_{\alpha,-}\,,\qquad J_- = J_x - iJ_y = \hbar\sum_\alpha A^\dagger_{\alpha,-}\,A_{\alpha,+}\,. \tag{G.72}$$

It is easy to show that J_i obeys the angular momentum algebra:

$$[\,J_i, J_j\,] = i\hbar\,\epsilon_{ijk}\,J_k\,, \tag{G.73}$$

or

$$[\,J_z, J_\pm\,] = \pm\hbar\,J_\pm\,,\qquad [\,J_+, J_-\,] = 2\hbar\,J_z\,. \tag{G.74}$$

The square of the total angular momentum operator is:

$$\begin{aligned} J^2 &= J_x^2 + J_y^2 + J_z^2 = \frac{1}{2}\,(\,J_+J_- + J_-J_+\,) + J_z^2 = J_+J_- + J_z^2 - \hbar\,J_z = J_-J_+ + J_z^2 + \hbar\,J_z\,,\\ &= \Big(\frac{\hbar}{2}\Big)^2\sum_{\alpha,\beta}\sum_{m,m',m'',m'''} A^\dagger_{\alpha,m}\,A_{\alpha,m'}A^\dagger_{\beta,m''}\,A_{\beta,m'''}\,\big(\,\boldsymbol{\sigma}_{m,m'}\cdot\boldsymbol{\sigma}_{m'',m'''}\,\big)\\ &= \Big(\frac{\hbar}{2}\Big)^2\sum_{\alpha,\beta}\sum_{m,m'}\big\{\,2\,A^\dagger_{\alpha,m}\,A_{\alpha,m'}A^\dagger_{\beta,m'}\,A_{\beta,m} - N_{\alpha,m}\,N_{\beta,m'}\,\big\}\,. \end{aligned} \tag{G.75}$$

In the last line, we have used the identity:

$$\big(\,\boldsymbol{\sigma}_{m,m'}\cdot\boldsymbol{\sigma}_{m'',m'''}\,\big) = 2\,\delta_{m,m'''}\,\delta_{m',m''} - \delta_{m,m'}\,\delta_{m'',m'''}\,. \tag{G.76}$$

So eigenvectors of J^2 and J_z obey the equations:

$$\begin{aligned} J^2\,|\,j,m\,) &= \hbar^2\,j(j+1)\,|\,j,m\,)\,,\\ J_z\,|\,j,m\,) &= \hbar\,m\,|\,j,m\,)\,,\\ J_\pm\,|\,j,m\,) &= \hbar\,A(j,\mp m)\,|\,j,m\pm1\,)\,, \end{aligned} \tag{G.77}$$

with $j = 0, 1/2, 1, 3/2, \ldots$, $-j \le m \le +j$, and $A(j,m) = \sqrt{(j+m)(j-m+1)}$.

We also define a Hermitian vector operator $\mathbf{K}$ by the mapping:

$$\mathbf{K} = \frac{\hbar}{2}\,\mathrm{Tr}[\,\boldsymbol{\sigma}^T\,A^\dagger\,A\,] = \frac{\hbar}{2}\sum_{\alpha,\beta,m}\boldsymbol{\sigma}_{\beta,\alpha}\,A^\dagger_{\alpha,m}\,A_{\beta,m} = \mathbf{K}_+ + \mathbf{K}_-\,. \tag{G.78}$$

Explicitly, we find:

$$\begin{aligned} K_x &= \frac{\hbar}{2}\sum_m \left(A^\dagger_{2,m}\, A_{1,m} + A^\dagger_{1,m}\, A_{2,m} \right) , \\ K_y &= \frac{\hbar}{2i}\sum_m \left(A^\dagger_{2,m}\, A_{1,m} - A^\dagger_{1,m}\, A_{2,m} \right) , \\ K_z &= \frac{\hbar}{2}\sum_m \left(A^\dagger_{1,m}\, A_{1,m} - A^\dagger_{2,m}\, A_{2,m} \right) = \frac{\hbar}{2}\left(N_1 - N_2 \right) , \end{aligned} \tag{G.79}$$

and

$$K_+ = K_x + iK_y = \hbar\sum_m A^\dagger_{2,m}\, A_{1,m} , \qquad K_- = K_x - iK_y = \hbar\sum_m A^\dagger_{1,m}\, A_{2,m} . \tag{G.80}$$

It is easy to show that K_i obeys the (hyperbolic) angular momentum algebra with a negative sign:

$$[\, K_i, K_j \,] = -i\hbar\, \epsilon_{ijk}\, K_k . \tag{G.81}$$

or

$$[\, K_z, K_\pm \,] = \mp\hbar\, K_\pm , \qquad [\, K_+, K_- \,] = -2\hbar\, K_z . \tag{G.82}$$

The square of the $\mathbf{K}$ vector is:

$$\begin{aligned} K^2 = K_x^2 + K_y^2 + K_z^2 &= \frac{1}{2}\left(K_+K_- + K_-K_+ \right) + K_z^2 = K_+K_- + K_z^2 + \hbar\, K_z = K_-K_+ + K_z^2 - \hbar\, K_z , \\ &= \left(\frac{\hbar}{2}\right)^2 \sum_{m,m'} \sum_{\alpha,\beta,\alpha',\beta'} A^\dagger_{\alpha,m}\, A_{\beta,m}\, A^\dagger_{\alpha',m'}\, A_{\beta',m'} \left(\boldsymbol{\sigma}_{\beta,\alpha} \cdot \boldsymbol{\sigma}_{\beta',\alpha'} \right) \\ &= \left(\frac{\hbar}{2}\right)^2 \sum_{m,m'} \sum_{\alpha,\beta} \left\{ 2\, A^\dagger_{\alpha,m}\, A_{\beta,m}\, A^\dagger_{\beta,m'}\, A_{\alpha,m'} - N_{\alpha,m}\, N_{\beta,m'} \right\} . \end{aligned} \tag{G.83}$$

Theorem 73. *Common eigenvectors of K^2 and K_z are:*

$$\begin{aligned} K^2 \,|\, k, q \,) &= \hbar^2\, k(k+1) \,|\, k, q \,) , \\ K_z \,|\, k, q \,) &= \hbar\, q \,|\, k, q \,) , \\ K_\pm \,|\, k, q \,) &= \hbar\, A(k, \pm q) \,|\, k, q \mp 1 \,) , \end{aligned} \tag{G.84}$$

with $k = 0, 1/2, 1, 3/2, \ldots$, $-k \le q \le +k$, and $A(k,q) = \sqrt{(k+q)(k-q+1)}$. These eigenvectors are similar to the eigenvectors of J^2 and J_z in Eqs. (G.77), except that the role of $K_\pm$ is reversed; K_+ on these eigenvectors decreases *the q-value by one, and K_-* increases *the q-value by one.*

Proof. The proof is left as an exercise. □

Theorem 74. *For* $\mathbf{J}$*, defined in Eq. (G.70), and* $\mathbf{K}$*, defined in Eq. (G.78), the square of the vector operators J^2 and K^2 are equal: $J^2 = K^2$.*

Proof. Starting with the last line of Eq. (G.83), we have:

$$\begin{aligned}
K^2 &= \Big(\frac{\hbar}{2}\Big)^2 \sum_{m,m'}\sum_{\alpha,\beta}\{\,2\,A^\dagger_{\alpha,m}\,A_{\beta,m}\,A^\dagger_{\beta,m'}\,A_{\alpha,m'} - N_{\alpha,m}\,N_{\beta,m'}\,\}\\
&= \Big(\frac{\hbar}{2}\Big)^2 \sum_{m,m'}\sum_{\alpha,\beta}\{\,2\,A^\dagger_{\alpha,m}\,A_{\beta,m}\,\{\,A_{\alpha,m'}\,A^\dagger_{\beta,m'} - [\,A_{\alpha,m'},A^\dagger_{\beta,m'}\,]\,\} - N_{\alpha,m}\,N_{\beta,m'}\,\}\\
&= \Big(\frac{\hbar}{2}\Big)^2 \sum_{m,m'}\sum_{\alpha,\beta}\{\,2\,\{\,A^\dagger_{\alpha,m}\,A_{\alpha,m'}\,A_{\beta,m}\,A^\dagger_{\beta,m'} - A^\dagger_{\alpha,m}\,A_{\alpha,m}\,\} - N_{\alpha,m}\,N_{\beta,m'}\,\}\\
&= \Big(\frac{\hbar}{2}\Big)^2 \sum_{m,m'}\sum_{\alpha,\beta}\{\,2\,\{\,A^\dagger_{\alpha,m}\,A_{\alpha,m'}\,\{\,A^\dagger_{\beta,m'}\,A_{\beta,m} + [\,A_{\beta,m},A^\dagger_{\beta,m'}\,]\,\} - A^\dagger_{\alpha,m}\,A_{\alpha,m}\,\} - N_{\alpha,m}\,N_{\beta,m'}\,\}\\
&= \Big(\frac{\hbar}{2}\Big)^2 \sum_{m,m'}\sum_{\alpha,\beta}\{\,2\,\{\,A^\dagger_{\alpha,m}\,A_{\alpha,m'}\,A^\dagger_{\beta,m'}\,A_{\beta,m} + A^\dagger_{\alpha,m}\,A_{\alpha,m} - A^\dagger_{\alpha,m}\,A_{\alpha,m}\,\} - N_{\alpha,m}\,N_{\beta,m'}\,\}\\
&= \Big(\frac{\hbar}{2}\Big)^2 \sum_{m,m'}\sum_{\alpha,\beta}\{\,2\,A^\dagger_{\alpha,m}\,A_{\alpha,m'}\,A^\dagger_{\beta,m'}\,A_{\beta,m} - N_{\alpha,m}\,N_{\beta,m'}\,\}\,.
\end{aligned} \tag{G.85}$$

which agrees with the last line of Eq. (G.75). So $K^2 = J^2$, which is what we were trying to prove. □

It is also easy to show that J_i commutes with all components of K_j:

$$[\,J_i,K_j\,] = \sum_{m,m',\alpha}\sum_{\alpha',\beta',m''}\sigma_{(i)\,m,m'}\,\sigma_{(j)\,\beta',\alpha'}\,[\,A^\dagger_{\alpha,m}\,A_{\alpha,m'},A^\dagger_{\alpha',m''}\,A_{\beta',m''}\,] = 0\,. \tag{G.86}$$

The Boson number operator S is defined by:

$$S = \mathrm{Tr}[\,A^\dagger A\,] = \sum_{\alpha,m} A^\dagger_{\alpha,m}\,A_{\alpha,m} = N_{1,+} + N_{1,-} + N_{2,+} + N_{2,-}\,, \tag{G.87}$$

which commutes with all operators. So we can find common eigenvectors of $J^2 = K^2$, J_z, K_z, and S. These eigenvectors are defined by:

$$J^2\,|\,j,m,q,s\,) = \hbar^2\,j(j+1)\,|\,j,m,q,s\,)\,, \qquad j = 0,1/2,1,3/2,2,\ldots \tag{G.88}$$
$$J_z\,|\,j,m,q,s\,) = \hbar\,m\,|\,j,m,q,s\,)\,, \qquad -j \le m \le +j\,, \tag{G.89}$$
$$K_z\,|\,j,m,q,s\,) = \hbar\,q\,|\,j,m,q,s\,)\,, \qquad -j \le q \le +j\,, \tag{G.90}$$
$$S\,|\,j,m,q,s\,) = 2s\,|\,j,m,q,s\,)\,, \qquad s = 0,1/2,1,3/2,2,\ldots \tag{G.91}$$

Note that s has half-integer values. From our previous results, we know that:

$$j_1 = (\,n_{1,+} + n_{1,-}\,)/2\,, \qquad m_1 = (\,n_{1,+} - n_{1,-}\,)/2\,,$$
$$j_2 = (\,n_{2,+} + n_{2,-}\,)/2\,, \qquad m_2 = (\,n_{2,+} - n_{2,-}\,)/2\,.$$

From the above, we also have:

$$m = \frac{1}{2}\,(\,n_{1,+} + n_{2,+} - n_{1,-} - n_{2,-}\,) = m_1 + m_2\,, \tag{G.92}$$

and

$$q = \frac{1}{2}\,(\,n_{1,+} + n_{1,-} - n_{2,+} - n_{2,-}\,) = j_1 - j_2\,. \tag{G.93}$$

From Eq. (G.90), we find the triangle inequality: $|j_1 - j_2| \leq j$. We also see that:

$$s = \frac{1}{2}\left(n_{1,+} + n_{1,-} + n_{2,+} + n_{2,-} \right) = j_1 + j_2 \,. \tag{G.94}$$

That is $j_1 = (s+q)/2$ and $j_2 = (s-q)/2$. So instead of labeling the vectors by (q,s) we can use the set (j_1, j_2), and write:

$$|\, j, m, q, s \,) \mapsto |\, (j_1, j_2)\, j, m \,) \,. \tag{G.95}$$

Now we want to find the states $|\, j, m, q, s \,\rangle$. We state the result in the form of the following theorem.

Theorem 75. *The coupled state* $|\, j, m, q, s \,)$ *is given by:*

$$|\, j, m, q, s \,) = \sqrt{\frac{(2j+1)}{(s-j)!\,(s+j+1)!}}\, \left[\det[\, A^\dagger \,] \right]^{s-j} D^{(j)\,\dagger}_{m,q}(\, A \,)\, |\, 0 \,) \,, \tag{G.96}$$

where $D_{m,q}(\, A \,)$ *is the D-matrix given in Eq. (G.55), with* $D_{+,+}(R) = A_{1,+}$, $D_{+,-}(R) = A_{2,+}$, $D_{-,+}(R) = A_{1,-}$, *and* $D_{-,-}(R) = A_{2,-}$.

Proof. We follow a method due to Sharp [5] here. We start by constructing the top state $|\, j, j, j, s \,)$ with $m = j$ and $q = j$. This state is defined by:

$$J_+ \,|\, j, j, j, s \,) = \hbar \left\{ A^\dagger_{1,+} A_{1,-} + A^\dagger_{2,+} A_{2,-} \right\} |\, j, j, j, s \,) = 0 \,, \tag{G.97}$$

$$K_- \,|\, j, j, j, s \,) = \hbar \left\{ A^\dagger_{1,+} A_{2,+} + A^\dagger_{1,-} A_{2,-} \right\} |\, j, j, j, s \,) = 0 \,. \tag{G.98}$$

We first note that:

$$\det[\, A^\dagger \,] = A^\dagger_{1,+} A^\dagger_{2,-} - A^\dagger_{1,-} A^\dagger_{2,+} = [\, \det[\, A \,] \,]^\dagger \,, \tag{G.99}$$

and that J_+ and K_- commute with $\det[\, A^\dagger \,]$:

$$[\, J_+, \det[\, A^\dagger \,] \,] = [\, K_-, \det[\, A^\dagger \,] \,] = 0 \,. \tag{G.100}$$

So in order to satisfy (G.97), $|\, j, j, j, s \,)$ must be of the general form:

$$|\, j, j, j, s \,) = \sum_{\alpha,\beta,\gamma} C_{\alpha,\beta,\gamma} \left[\det[\, A^\dagger \,] \right]^\alpha \left[A^\dagger_{1,+} \right]^\beta \left[A^\dagger_{2,+} \right]^\gamma |\, 0 \,) \,. \tag{G.101}$$

Now K_- commutes with $A^\dagger_{1,+}$ but not with $A^\dagger_{2,+}$, so we must have $C_{\alpha,\beta,\gamma} = \delta_{\gamma,0}\, C_{\alpha,\beta}$. So in order to satisfy (G.98), $|\, j, j, j, s \,)$ must be of the general form:

$$|\, j, j, j, s \,) = \sum_{\alpha,\beta} C_{\alpha,\beta} \left[\det[\, A^\dagger \,] \right]^\alpha \left[A^\dagger_{1,+} \right]^\beta |\, 0 \,) \,. \tag{G.102}$$

In addition, since J_z and K_z also commute with $\det[\, A^\dagger \,]$:

$$[\, J_z, \det[\, A^\dagger \,] \,] = [\, K_z, \det[\, A^\dagger \,] \,] = 0 \,, \tag{G.103}$$

we find that:

$$\begin{aligned} J_z \,|\, j, j, j, s \,) &= \sum_{\alpha,\beta} C_{\alpha,\beta} \left[\det[\, A^\dagger \,] \right]^\alpha J_z \left[A^\dagger_{1,+} \right]^\beta |\, 0 \,) = \frac{\hbar}{2} \sum_{\alpha,\beta} \beta\, C_{\alpha,\beta} \left[\det[\, A^\dagger \,] \right]^\alpha \left[A^\dagger_{1,+} \right]^\beta |\, 0 \,) \\ &= \hbar\, j \,|\, j, j, j, s \,) \,, \end{aligned} \tag{G.104}$$

so $C_{\alpha,\beta} = \delta_{\beta,2j}\, C_\alpha$. This also works for K_z, as can be easily checked. So we conclude that:

$$|\, j, j, j, s \,) = \sum_{\alpha} C_\alpha \left[\det[\, A^\dagger \,] \right]^\alpha \left[A^\dagger_{1,+} \right]^{2j} |\, 0 \,) \,. \tag{G.105}$$

Now since S is the Bose number operator, by Euler's theorem on homogeneous functions, Theorem 61 in Appendix B, the eigenvector $|\,j,j,j,s\,\rangle$ must be a homogeneous function of the creation operators of degree $2s$. That is, since $[\,S,\det[\,A^\dagger\,]^\alpha\,]=2\alpha\,\det[\,A^\dagger\,]^\alpha$, we find:

$$S\,|\,j,j,j,s\,)=\sum_\alpha C_\alpha\,(\,2\alpha+2j\,)\,\big[\,\det[\,A^\dagger\,]\,\big]^\alpha\,\big[\,A^\dagger_{1,+}\,\big]^{2j}\,|\,0\,)=2s\,|\,j,j,j,s\,)\,, \tag{G.106}$$

so we conclude that $C_\alpha=\delta_{\alpha,s-j}\,C$. Then the top eigenvector is given by:

$$\begin{aligned}
|\,j,j,j,s\,) &= C\,\big[\,\det[\,A^\dagger\,]\,\big]^{s-j}\,\big[\,A^\dagger_{1,+}\,\big]^{2j}\,|\,0\,)\\
&= C\sum_{n=0}^{s-j}(-)^n\,\binom{s-j}{n}\,(\,A^\dagger_{1,+}A^\dagger_{2,-}\,)^{s-j-n}\,(\,A^\dagger_{1,-}A^\dagger_{2,+}\,)^n\,\big[\,A^\dagger_{1,+}\,\big]^{2j}\,|\,0\,)\\
&= C\sum_{n=0}^{s-j}(-)^n\,\binom{s-j}{n}\,n!\,\sqrt{(s+j-n)!\,(s-j-n)!}\,|\,s+j-n,n,n,s-j-n\,)\\
&= C\,(s-j)!\sum_{n=0}^{s-j}(-)^n\,\sqrt{\frac{(s+j-n)!}{(s-j-n)!}}\,|\,s+j-n,n,n,s-j-n\,)\,.
\end{aligned} \tag{G.107}$$

The normalization requirement fixes the value of C. That is:

$$|\,C\,|^2\,[\,(s-j)!\,]^2\sum_{n=0}^{s-j}\frac{(s+j-n)!}{(s-j-n)!}=|\,C\,|^2\,\frac{(s-j)!\,(s+j+1)!}{(2j+1)}=1\,, \tag{G.108}$$

where we have used Eq. (C.12) in Appendix C. So we find that:

$$C=\sqrt{\frac{(2j+1)}{(s-j)!\,(s+j+1)!}}\,. \tag{G.109}$$

The phase is arbitrary, and chosen here to be one, which will be explained later. So from (G.107), the vector $|\,j,j,j,s\,)$ is given by:

$$|\,j,j,j,s\,)=\sqrt{\frac{(2j+1)}{(s-j)!\,(s+j+1)!}}\,\big[\,\det[\,A^\dagger\,]\,\big]^{s-j}\,\big[\,A^\dagger_{1,+}\,\big]^{2j}\,|\,0\,)\,. \tag{G.110}$$

The vector $|\,j,j,q,s\,)$ is obtained by operating $j-q$ times by K_+ on (G.110):

$$
\begin{aligned}
|\,j,j,q,s\,) &= \frac{1}{\hbar^{j-q}}\sqrt{\frac{(j+q)!}{(2j)!\,(j-q)!}}\,[\,K_+\,]^{j-q}\,|\,j,j,j,s\,) \\
&= \sqrt{\frac{(2j+1)\,(j+q)!}{(s-j)!\,(s+j+1)!\,(2j)!\,(j-q)!}}\,(\,A^\dagger_{2,+}\,A_{1,+}+A^\dagger_{2,-}\,A_{1,-}\,)^{j-q}\,[\,\det[\,A^\dagger\,]\,]^{s-j}\,[\,A^\dagger_{1,+}\,]^{2j}\,|\,0\,) \\
&= \sqrt{\frac{(2j+1)\,(j+q)!}{(s-j)!\,(s+j+1)!\,(2j)!\,(j-q)!}}\,[\,\det[\,A^\dagger\,]\,]^{s-j} \\
&\qquad\times\sum_{n=0}^{j-q}\binom{j-q}{n}\,[\,A^\dagger_{2,+}\,A_{1,+}\,]^{j-q-n}\,[\,A^\dagger_{1,+}\,]^{2j}\,[\,A^\dagger_{2,-}\,A_{1,-}\,]^{n}\,|\,0\,) \\
&= \sqrt{\frac{(2j+1)\,(j+q)!}{(s-j)!\,(s+j+1)!\,(2j)!\,(j-q)!}}\,[\,\det[\,A^\dagger\,]\,]^{s-j}\,[\,A^\dagger_{2,+}\,A_{1,+}\,]^{j-q}\,[\,A^\dagger_{1,+}\,]^{2j}\,|\,0\,) \\
&= \sqrt{\frac{(2j+1)\,(j+q)!}{(s-j)!\,(s+j+1)!\,(2j)!\,(j-q)!}}\,[\,\det[\,A^\dagger\,]\,]^{s-j}\,[\,A^\dagger_{2,+}\,]^{j-q}\,[\,(\,A_{1,+}\,)^{j-q},(\,A^\dagger_{1,+}\,)^{2j}\,]\,|\,0\,) \\
&= \sqrt{\frac{(2j+1)!}{(s-j)!\,(s+j+1)!\,(j+q)!\,(j-q)!}}\,[\,\det[\,A^\dagger\,]\,]^{s-j}\,[\,A^\dagger_{2,+}\,]^{j-q}\,[\,A^\dagger_{1,+}\,]^{j+q}\,|\,0\,)
\end{aligned}
\tag{G.111}
$$

Here, we have used the fact that K_+ commutes with $\det[\,A^\dagger\,]$ and Eq. (B.3). Finally, the vector $|\,j,m,q,s\,)$ is found by operating $j-m$ times by J_- on (G.111):

$$
\begin{aligned}
|\,j,m,q,s\,) &= \frac{1}{\hbar^{j-m}}\sqrt{\frac{(j+m)!}{(2j)!\,(j-m)!}}\,[\,J_-\,]^{j-m}\,|\,j,j,q,s\,\rangle \\
&= \sqrt{\frac{(2j+1)\,(j+m)!}{(j-m)!\,(s-j)!\,(s+j+1)!\,(j+q)!\,(j-q)!}}\,(\,A^\dagger_{1,-}\,A_{1,+}+A^\dagger_{2,-}\,A_{2,+}\,)^{j-m} \\
&\qquad\times[\,\det[\,A^\dagger\,]\,]^{s-j}\,[\,A^\dagger_{2,+}\,]^{j-q}\,[\,A^\dagger_{1,+}\,]^{j+q}\,|\,0\,)\,.
\end{aligned}
\tag{G.112}
$$

Now using the fact that J_- commutes with $\det[\,A^\dagger\,]$, Eq. (G.112) becomes:

$$
\begin{aligned}
|\,j,m,q,s\,) &= \sqrt{\frac{(2j+1)\,(j+m)!}{(j-m)!\,(s-j)!\,(s+j+1)!\,(j+q)!\,(j-q)!}}\,[\,\det[\,A^\dagger\,]\,]^{s-j} \\
&\qquad\times\sum_{n=0}^{j-m}\binom{j-m}{n}\,[\,A^\dagger_{1,-}\,A_{1,+}\,]^{n}\,[\,A^\dagger_{1,+}\,]^{j+q}\,[\,A^\dagger_{2,-}\,A_{2,+}\,]^{j-m-n}\,[\,A^\dagger_{2,+}\,]^{j-q}\,|\,0\,) \\
&= \sqrt{\frac{(2j+1)\,(j+m)!}{(j-m)!\,(s-j)!\,(s+j+1)!\,(j+q)!\,(j-q)!}}\,[\,\det[\,A^\dagger\,]\,]^{s-j} \\
&\qquad\times\sum_{n=0}^{j-m}\binom{j-m}{n}\,[\,(\,A_{1,+}\,)^{n},(\,A^\dagger_{1,+}\,)^{j+q}\,][\,(\,A_{2,+}\,)^{j-m-n},(\,A^\dagger_{2,+}\,)^{j-q}\,] \\
&\qquad\qquad\times(\,A^\dagger_{1,-}\,)^{n}\,(\,A^\dagger_{2,-}\,)^{j-m-n}\,|\,0\,)\,.
\end{aligned}
\tag{G.113}
$$

Again, using (B.3), Eq. (G.113) becomes:

$$
\begin{aligned}
|\, j,m,q,s\,) &= \sqrt{\frac{(2j+1)\,(j+m)!\,(j-m)!\,(j+q)!\,(j-q)!}{(s-j)!\,(s+j+1)!}}\,\big[\det[\,A^\dagger\,]\,\big]^{s-j} \\
&\qquad \times \sum_{n=0}^{j-m} \frac{\big(\,A^\dagger_{1,+}\,\big)^{j+q-n}\,\big(\,A^\dagger_{2,+}\,\big)^{m-q+n}\,\big(\,A^\dagger_{1,-}\,\big)^{n}\,\big(\,A^\dagger_{2,-}\,\big)^{j-m-n}}{(j+q-n)!\,(m-q+n)!\,(n)!\,(j-m-n)!}\,|\,0\,) \\
&= \sqrt{\frac{(2j+1)\,(j+m)!\,(j-m)!\,(j+q)!\,(j-q)!}{(s-j)!\,(s+j+1)!}}\,\big[\det[\,A^\dagger\,]\,\big]^{s-j} \\
&\qquad \times \sum_{n=0}^{j-m}\sum_{n'=0}^{j+m} \delta_{n'-n,m-q}\frac{\big(\,A^\dagger_{1,+}\,\big)^{j+m-n'}\,\big(\,A^\dagger_{2,+}\,\big)^{n'}\,\big(\,A^\dagger_{1,-}\,\big)^{n}\,\big(\,A^\dagger_{2,-}\,\big)^{j-m-n}}{(j+m-n')!\,(n')!\,(n)!\,(j-m-n)!}\,|\,0\,)\,, \\
&= \sqrt{\frac{(2j+1)}{(s-j)!\,(s+j+1)!}}\,\big[\det[\,A^\dagger\,]\,\big]^{s-j}\,D^{(j)\dagger}_{m,q}(\,A\,)\,|\,0\,)\,,
\end{aligned}
\tag{G.114}
$$

where $D_{m,q}(\,A\,)$ is the D-matrix given in Eq. (G.55), with $D_{+,+}(R) = A_{1,+}$, $D_{+,-}(R) = A_{2,+}$, $D_{-,+}(R) = A_{1,-}$, and $D_{-,-}(R) = A_{2,-}$, which is what we were trying to prove. □

For our case, $s = j_1 + j_2$ and $q = j_1 - j_2$, so Theorem 75 states that the coupled angular momentum state $|\,(j_1,j_2)\,j,m\,)$ in the Bose representation is given by:

$$
|\,(j_1,j_2)\,j,m\,) = \sqrt{\frac{(2j+1)}{(j_1+j_2-j)!\,(j_1+j_2+j+1)!}}\,\Big[\,\big[\det[\,A\,]\,\big]^{j_1+j_2-j}\,D^{(j)}_{m,j_1-j_2}(\,A\,)\,\Big]^\dagger\,|\,0\,)\,. \tag{G.115}
$$

This formula was first stated by Biedenharn [3][p. 225]. We also know from Eq. (G.12) that the uncoupled vector $|\,j_1,m_1,j_2,m_2\,)$ is given by:

$$
|\,j_1,m_1,j_2,m_2\,) = \frac{\big(\,A^\dagger_{1,+}\,\big)^{j_1+m_1}\,\big(\,A^\dagger_{1,-}\,\big)^{j_1-m_1}\,\big(\,A^\dagger_{2,+}\,\big)^{j_2+m_2}\,\big(\,A^\dagger_{2,-}\,\big)^{j_2-m_2}}{\sqrt{(j_1+m_1)!\,(j_1-m_1)!\,(j_2+m_2)!\,(j_2-m_2)!}}\,|\,0\,)\,. \tag{G.116}
$$

Clebsch-Gordan coefficients are the overlap between these two vectors. We can easily find a closed formula for these coefficients by expanding (G.115) and picking out the coefficients of $|\,j_1,m_1,j_2,m_2\,)$. This gives:

$$
\begin{aligned}
|\,(j_1,j_2)\,j,m\,) &= \sqrt{\frac{(2j+1)\,(j+m)!\,(j-m)!\,(j+j_1-j_2)!\,(j-j_1+j_2)!}{(j_1+j_2-j)!\,(j_1+j_2+j+1)!}}\,\big[\,A^\dagger_{1,+}\,A^\dagger_{2,-} - A^\dagger_{1,-}\,A^\dagger_{2,+}\,\big]^{j_1+j_2-j} \\
&\qquad \times \sum_{n'=0}^{j+m}\sum_{n=0}^{j-m} \delta_{n'-n,m-j_1+j_2}\frac{\big(\,A^\dagger_{1,+}\,\big)^{j+m-n'}\,\big(\,A^\dagger_{2,+}\,\big)^{n'}\,\big(\,A^\dagger_{1,-}\,\big)^{n}\,\big(\,A^\dagger_{2,-}\,\big)^{j-m-n}}{(j+m-n')!\,(n')!\,(n)!\,(j-m-n)!}\,|\,0\,)\,, \\
&= \sqrt{\frac{(2j+1)\,(j+m)!\,(j-m)!\,(j+j_1-j_2)!\,(j-j_1+j_2)!}{(j_1+j_2-j)!\,(j_1+j_2+j+1)!}} \\
&\qquad \times \sum_{n'=0}^{j+m}\sum_{n=0}^{j-m}\sum_{k=0}^{j_1+j_2-j}(-)^k\binom{j_1+j_2-j}{k}\,\delta_{n'-n,m-j_1+j_2} \\
&\qquad \times \frac{\big(\,A^\dagger_{1,+}\,\big)^{m-n'+j_1+j_2-k}\,\big(\,A^\dagger_{1,-}\,\big)^{n+k}\,\big(\,A^\dagger_{2,+}\,\big)^{n'+k}\,\big(\,A^\dagger_{2,-}\,\big)^{-m-n+j_1+j_2-k}}{(j+m-n')!\,(n')!\,(n)!\,(j-m-n)!}\,|\,0\,)\,.
\end{aligned}
\tag{G.117}
$$

So let us put $m_1 = m - n' + j_2 - k = j_1 - n - s$. So $n + k = j_1 - m_1$. We also put $n' + k = j_2 + m_2 = j_2 + m - m_1$. Then $-m - n + j_1 + j_2 - k = j_2 - m_2$. This means that we can set $n = j_1 - m_1 - k$ and $n' = j_2 + m_2 - k$,

so that (G.117) becomes:

$$|\,(j_1,j_2)\,j,m\,) = \sum_{m_1,m_2} \delta_{m,m_1+m_2} \sqrt{\frac{(2j+1)\,(j+m)!\,(j-m)!\,(j+j_1-j_2)!\,(j-j_1+j_2)!\,(j_1+j_2-j)!}{(j_1+j_2+j+1)!}}$$
$$\sum_{k=0}^{j_1+j_2-j} \frac{(-)^k\,(\,A^\dagger_{1,+}\,)^{j_1+m_1}\,(\,A^\dagger_{1,-}\,)^{j_1-m_1}\,(\,A^\dagger_{2,+}\,)^{j_2+m_2}\,(\,A^\dagger_{2,-}\,)^{j_2-m_2}|\,0\,)}{(j_1+j_2-j-k)!\,k!\,(j+m_1-j_2+k)!\,(j_2+m-m_1-k)!\,(j_1-m_1-k)!\,(j-m-j_1+m_1+k)!}$$
$$= \sum_{m_1,m_2} |\,j_1,m_1,j_2,m_2\,)\,\langle\,j_1,m_1,j_2,m_2\,|\,(j_1,j_2)\,j,m\,\rangle\,, \quad \text{(G.118)}$$

where the Clebsch-Gordan coefficient is given by:

$$\langle\,j_1,m_1,j_2,m_2\,|\,(j_1,j_2)\,j,m\,\rangle = \delta_{m,m_1+m_2}\,\sqrt{(j+m)!\,(j-m)!\,(j_1+m_1)!\,(j_1-m_1)!\,(j_2+m_2)!\,(j_2-m_2)!}$$
$$\times\sqrt{\frac{(2j+1)\,(j+j_1-j_2)!\,(j-j_1+j_2)!\,(j_1+j_2-j)!}{(j_1+j_2+j+1)!}}$$
$$\times\sum_k \frac{(-)^k}{k!\,(j_1+j_2-j-k)!\,(j-j_2+m_1+k)!\,(j_2+m_2-k)!\,(j_1-m_1-k)!\,(j-j_1-m_2+k)!}\,. \quad \text{(G.119)}$$

This is called "Racah's second form" for the Clebsch-Gordan coefficients. It can be shown to be identical to Eq. (21.196).

Exercise 92. Using relations in Appendix C, show that Eqs. (21.196) and (G.119) are identical (See Edmonds [4][p. 44–45]).

G.7 Generating function

Theorem 76 (Schwinger's generating function)**.** *A generating function $G(a,b)$ for the $3j$-symbols is:*

$$G(a,b) = \sum_{all\ j,\,m} f_{j_1,m_1}(a_1)\,f_{j_2,m_2}(a_2)\,f_{j_3,m_3}(a_3)\,F_{j_1,j_2,j_3}(b_1,b_2,b_3)\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$$
$$= \exp\{\,(\,a_1,a_2\,)\,b_3 + (\,a_2,a_3\,)\,b_1 + (\,a_3,a_1\,)\,b_2\,\}\,, \quad \text{(G.120)}$$

where $(a_i,a_j) := a_{i,+}\,a_{j,-} - a_{i,-}\,a_{j,+}$ and where $f_{j,m}(a)$ and $F_{j_1,j_2,j_3}(b_1,b_2,b_3)$ are given by:

$$f_{j,m}(a) = \frac{(\,a_+\,)^{j+m}\,(\,a_-\,)^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}\,, \quad \text{(G.121)}$$

$$F_{j_1,j_2,j_3}(b_1,b_2,b_3) = \sqrt{(j_1+j_2+j_3+1)!}\,\frac{(\,b_1\,)^{-j_1+j_2+j_3}\,(\,b_2\,)^{+j_1-j_2+j_3}\,(\,b_3\,)^{+j_1+j_2-j_3}}{\sqrt{(-j_1+j_2+j_3)!\,(j_1-j_2+j_3)!\,(j_1+j_2-j_3)!}}\,. \quad \text{(G.122)}$$

Proof. Following Schwinger [1], we first write the eigenvalue equation for the Bose operators as:

$$A_{\alpha,m}\,|\,a_{1,+},a_{1,-},a_{2,+},a_{2,-}\,) = a_{\alpha,m}\,|\,a_{1,+},a_{1,-},a_{2,+},a_{2,-}\,)\,, \quad \text{(G.123)}$$

where $a_{\alpha,m}$ is a complex number with $\alpha = 1,2$ and $m = \pm$. So from Eq. (G.116), the overlap of the coherent state with the uncoupled state is given by:

$$(\,j_1,m_1,j_2,m_2\,|\,a_{1,+},a_{1,-},a_{2,+},a_{2,-}\,) = \mathcal{N}(a_1)\,\mathcal{N}(a_2)\,f_{j_1,m_1}(a_1)\,f_{j_2,m_2}(a_2) \quad \text{(G.124)}$$

where

$$f_{j,m}(a) = \frac{(\,a_+\,)^{j+m}\,(\,a_-\,)^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}\,, \qquad \mathcal{N}(a) = \exp\Big\{-\sum_m a_m^*\,a_m/2\Big\}\,, \tag{G.125}$$

where we have normalized the coherent states according to Eq. (G.37). Then let us note that

$$\begin{aligned}\sum_{m_1,m_2} |\,j_1,m_1,j_2,m_2\,)\,f_{j_1,m_1}(a_1)\,f_{j_2,m_2}(a_2)\\ &= \sum_{m_1,m_2} \frac{\big(\,a_{1,+}\,A_{1,+}^\dagger\,\big)^{j_1+m_1}\,\big(\,a_{1,-}\,A_{1,-}^\dagger\,\big)^{j_1-m_1}\,\big(\,a_{2,+}\,A_{2,+}^\dagger\,\big)^{j_2+m_2}\,\big(\,a_{2,-}\,A_{2,-}^\dagger\,\big)^{j_2-m_2}}{(j_1+m_1)!\,(j_1-m_1)!\,(j_2+m_2)!\,(j_2-m_2)!}\,|\,0\,)\\ &= \frac{\big(\sum_m a_{1,m}\,A_{1,m}^\dagger\big)^{2j_1}\,\big(\sum_m a_{2,m}\,A_{2,m}^\dagger\big)^{2j_2}}{(2j_1)!\,(2j_2)!}\,|\,0\,)\end{aligned} \tag{G.126}$$

For the coupled state $|\,(j_1,j_2)\,j_3,m_3\,\rangle$ from (G.115), we have:

$$\begin{aligned}\sum_{m_3} |\,(j_1,j_2)\,j_3,-m_3\,)\,(-)^{j_1-j_2-m_3}\,f_{j_3,m_3}^*(a_3)/\sqrt{2j_3+1} &= \sqrt{\frac{(j_3+j_1-j_2)!\,(j_3-j_1+j_2)!}{(j_1+j_2-j_3)!\,(j_1+j_2+j_3+1)!}}\\ &\times \sum_{m_3}\sum_{n=0}^{j_3-m_3}\sum_{n'=0}^{j_3+m_3} \delta_{n'-n,-m_3-j_1+j_2}\,(-)^{j_1-j_2-j_3}\,\big(\,a_{3+}^*\,\big)^{j_3+m_3}\,\big(\,-a_{3-}^*\,\big)^{j_3-m_3}\\ &\times \frac{\big(\,A_{1,+}^\dagger\,\big)^{j_3-m_3-n'}\,\big(\,A_{2,+}^\dagger\,\big)^{n'}\,\big(\,A_{1,-}^\dagger\,\big)^{n}\,\big(\,A_{2,-}^\dagger\,\big)^{j_3+m_3-n}}{(j_3-m_3-n')!\,(n')!\,(n)!\,(j_3+m_3-n)!}\,\big(\,A_{1,+}^\dagger\,A_{2,-}^\dagger - A_{1,-}^\dagger\,A_{2,+}^\dagger\,\big)^{j_1+j_2-j_3}\,|\,0\,)\\ &= (-)^{-j_1+j_2+j_3}\sum_{n=0}^{j_3+j_1-j_2}\sum_{n'=0}^{j_3-j_1+j_2}\binom{j_3+j_1-j_2}{n}\binom{j_3-j_1+j_2}{n'}\,\big(\,A_{1,+}^\dagger\,A_{2,-}^\dagger - A_{1,-}^\dagger\,A_{2,+}^\dagger\,\big)^{j_1+j_2-j_3}\\ &\times \frac{\big(\,-a_{3,-}^*\,A_{1,+}^\dagger\,\big)^{j_3+j_1-j_2-n}\,\big(\,a_{3,+}^*\,A_{1,-}^\dagger\,\big)^{n}\,\big(\,a_{3,+}^*\,A_{2,-}^\dagger\,\big)^{j_3-j_1+j_2-n'}\,\big(\,-a_{3,-}^*\,A_{2,+}^\dagger\,\big)^{n'}}{\sqrt{(j_1+j_2+j_3+1)!\,(-j_1+j_2+j_3)!\,(j_1-j_2+j_3)!\,(j_1+j_2-j_3)!}}\,|\,0\,)\\ &= \frac{\big(\,-a_{3,+}^*\,A_{2,-}^\dagger + a_{3,-}^*\,A_{2,+}^\dagger\,\big)^{-j_1+j_2+j_3}\,\big(\,a_{3,+}^*\,A_{1,-}^\dagger - a_{3,-}^*\,A_{1,+}^\dagger\,\big)^{j_1-j_2+j_3}}{\sqrt{(j_1+j_2+j_3+1)!\,(-j_1+j_2+j_3)!\,(j_1-j_2+j_3)!}}\\ &\times \frac{\big(\,A_{1,+}^\dagger\,A_{2,-}^\dagger - A_{1,-}^\dagger\,A_{2,+}^\dagger\,\big)^{j_1+j_2-j_3}}{\sqrt{(j_1+j_2-j_3)!}}\,|\,0\,)\,.\end{aligned} \tag{G.127}$$

The overlap between Eqs. (G.126) and (G.127) is:

$$\begin{aligned}\sum_{m_1,m_2,m_3} f_{j_1,m_1}(a_1)\,f_{j_2,m_2}(a_2)\,f_{j_3,m_3}(a_3)\begin{pmatrix} j_1 & j_2 & j_3\\ m_1 & m_2 & m_3\end{pmatrix} &= \frac{1}{\sqrt{(j_1+j_2+j_3+1)!}}\\ \times\,(\,0\,|\,\frac{\big(\,A_{1,+}\,A_{2,-} - A_{1,-}\,A_{2,+}\,\big)^{j_1+j_2-j_3}}{\sqrt{(j_1+j_2-j_3)!}}\,\frac{\big(\,a_{3,+}\,A_{1,-} - a_{3,-}\,A_{1,+}\,\big)^{j_1-j_2+j_3}}{\sqrt{(j_1-j_2+j_3)!}}\,\frac{\big(\,a_{1,+}\,A_{1,+}^\dagger + a_{1,-}\,A_{1,-}^\dagger\,\big)^{2j_1}}{(2j_1)!}\\ \times\,\frac{\big(\,-a_{3,+}\,A_{2,-} + a_{3,-}\,A_{2,+}\,\big)^{-j_1+j_2+j_3}}{\sqrt{(-j_1+j_2+j_3)!}}\,\frac{\big(\,a_{2,+}\,A_{2,+}^\dagger + a_{2,-}\,A_{2,-}^\dagger\,\big)^{2j_2}}{(2j_2)!}\,|\,0\,)\,.\end{aligned} \tag{G.128}$$

Here, we want to move the creation operators to the left and the annihilation operators to the right. Using

Eq. (B.3), we find:

$$[\,(-a_{3,+}\,A_{2,-}+a_{3,-}\,A_{2,+}\,)^{-j_1+j_2+j_3},(\,a_{2,+}\,A^\dagger_{2,+}+a_{2,-}\,A^\dagger_{2,-}\,)^{2j_2}\,]\,|\,0\,)$$
$$=\frac{(2j_2)!}{(j_1+j_2-j_3)!}\,(\,-a_{3,+}\,a_{2,-}+a_{3,-}\,a_{2,+}\,)^{-j_1+j_2+j_3}\,(\,a_{2,+}\,A^\dagger_{2,+}+a_{2,-}\,A^\dagger_{2,-}\,)^{j_1+j_2-j_3}\,|\,0\,)\,, \quad \text{(G.129)}$$

and

$$[\,(a_{3,+}\,A_{1,-}-a_{3,-}\,A_{1,+}\,)^{j_1-j_2+j_3},(\,a_{1,+}\,A^\dagger_{1,+}+a_{1,-}\,A^\dagger_{1,-}\,)^{2j_1}\,]\,|\,0\,)$$
$$=\frac{(2j_1)!}{(j_1+j_2-j_3)!}\,(\,a_{3,+}\,a_{1,-}-a_{3,-}\,a_{1,+}\,)^{j_1-j_2+j_3}\,(\,a_{1,+}\,A^\dagger_{1,+}+a_{1,-}\,A^\dagger_{1,-}\,)^{j_1+j_2-j_3}\,|\,0\,)\,. \quad \text{(G.130)}$$

Now let us define D by:

$$D=A_{1,+}\,A_{2,-}-A_{1,-}\,A_{2,+}\,, \quad \text{(G.131)}$$

and call C_α:

$$C_\alpha=a_{\alpha,+}\,A^\dagger_{\alpha,+}+a_{\alpha,-}\,A^\dagger_{\alpha,-}\,,\qquad \alpha=1,2. \quad \text{(G.132)}$$

then the remaining term we need to calculate is:

$$(\,0\,|\,[\,D^{j_1+j_2-j_3},(\,C_1C_2\,)^{j_1+j_2-j_3}\,]\,|\,0\,)=(j_1+j_2-j_3)!\,\big[\,(\,0\,|\,[\,D,C_1C_2\,]\,|\,0\,)\,\big]^{j_1+j_2-j_3}\,, \quad \text{(G.133)}$$

since $D\,|\,0\,)=0$ and $(\,0\,|\,C_1=(\,0\,|\,C_2=0$. We find

$$\begin{aligned}[D,C_1]&=[\,A_{1,+},C_1\,]\,A_{2,-}-[\,A_{1,-},C_1\,]\,A_{2,+}\,,\\&=a_{1,+}\,A_{2,-}-a_{1,-}\,A_{2,+}\\[D,C_2]&=A_{1,+}[\,A_{2,-},C_2\,]-A_{1,-}[\,A_{2,+},C_2\,]\\&=a_{2,-}\,A_{1,+}-a_{2,+}\,A_{1,-}\,,\end{aligned} \quad \text{(G.134)}$$

so that

$$\begin{aligned}[D,C_1\,C_2\,]&=[\,D,C_1\,]\,C_2+C_1\,[\,D,C_2\,]\\&=(\,a_{1,+}\,A_{2,-}-a_{1,-}\,A_{2,+}\,)\,(\,a_{2,+}\,A^\dagger_{2,+}+a_{2,-}\,A^\dagger_{2,-}\,)\\&\quad+(\,a_{1,+}\,A^\dagger_{1,+}+a_{1,-}\,A^\dagger_{1,-}\,)\,(\,a_{2,-}\,A_{1,+}-a_{2,+}\,A_{1,-}\,)\,,\end{aligned} \quad \text{(G.135)}$$

so

$$(\,0\,|\,[\,D^{j_1+j_2-j_3},(\,C_1C_2\,)^{j_1+j_2-j_3}\,]\,|\,0\,)=(j_1+j_2-j_3)!\,(\,a_{1,+}\,a_{2,-}-a_{1,-}\,a_{2,+}\,)^{j_1+j_2-j_3}\,. \quad \text{(G.136)}$$

Then Eq. (G.128) becomes:[3]

$$\sum_{m_1,m_2,m_3}f_{j_1,m_1}(a_1)\,f_{j_2,m_2}(a_2)\,f_{j_3,m_3}(a_3)\begin{pmatrix}j_1&j_2&j_3\\m_1&m_2&m_3\end{pmatrix}=\frac{1}{\sqrt{(j_1+j_2+j_3+1)!}}$$
$$\times\frac{(\,a_{1,+}\,a_{2,-}-a_{1,-}\,a_{2,+}\,)^{j_1+j_2-j_3}\,(\,a_{2,+}\,a_{3,-}-a_{2,-}\,a_{3,+}\,)^{-j_1+j_2+j_3}\,(\,a_{3,+}\,a_{1,-}-a_{3,-}\,a_{1,+}\,)^{j_1-j_2+j_3}}{\sqrt{(j_1+j_2-j_3)!\,(-j_1+j_2+j_3)!\,(j_1-j_2+j_3)!}}\,. \quad \text{(G.137)}$$

Multiplying this equation on both sides by $F_{j_1,j_2,j_3}(b_1,b_2,b_3)$ and summing over j_1, j_2, and j_3 gives the result quoted for the generator in Eq. (G.120), which was what we were trying to prove. □

Theorem 76 provides an easy way to find all the symmetry properties of the $3j$-symbols.

[3] Help! There seems to be a extra factor of $(j_1+j_2-j_3)!$ left over. What happened?

G.8 Bose tensor operators

We have already seen one example of a tensor operator in our calculation of the D-functions using Bose operators. We showed in Eq. (G.52) that the Bose operators A_q are tensor operators of rank $k = 1/2$ and transform according to:

$$U^\dagger(R)\, A_q\, U(R) = \sum_{q'=-1/2}^{+1/2} D^{(1/2)}_{q,q'}(R)\, A_{q'}\,, \quad \text{and} \quad U^\dagger(R)\, A^\dagger_q\, U(R) = \sum_{q'=-1/2}^{+1/2} A^\dagger_{q'}\, D^{(1/2)}_{q',q}(R^{-1})\,. \tag{G.138}$$

So $T(1/2, q) = A^\dagger_q$ is a tensor operator. A_q transforms as an adjoint tensor operator. It is *not* a Hermitian tensor operator as defined in either Eq. (21.249) or (21.250). The number operator $N = A^\dagger A = \sum_q A^\dagger_q A_q$ is an invariant under rotations:

$$\begin{aligned} U^\dagger(R)\, N\, U(R) &= U^\dagger(R) \Big\{ \sum_q A^\dagger_q\, A_q \Big\} U(R) \\ &= \sum_{q,q',q''} A^\dagger_{q'}\, D^{(1/2)}_{q',q}(R^{-1})\, D^{(1/2)}_{q,q''}(R)\, A_{q''} = \sum_{q'} A^\dagger_{q'}\, A_{q'} = N\,. \end{aligned} \tag{G.139}$$

However the more general tensor product $N(S, M)$ defined by:

$$N(S, M) = \sum_{m,m'} \langle\, 1/2, m, 1/2, m' \,|\, (1/2, 1/2)\, S, M \,\rangle\, A^\dagger_m\, A_{m'}\,, \tag{G.140}$$

and transforms in a different way. [Work this out...]

We can construct tensor operators for the coupling of two commuting angular momenta also. Following our definitions in Section G.6, let $A^\dagger_{1,m}$ and $A^\dagger_{2,m}$ be two commuting sets of creation operators, $m = \pm$, obeying the algebra:

$$[\, A_{\alpha,m}, A^\dagger_{\beta,m'}\,] = \delta_{m,m'}\delta_{\alpha,\beta}\,, \qquad [\, A_{\alpha,m}, A_{\beta,m'}\,] = [\, A^\dagger_{\alpha,m}, A^\dagger_{\beta,m'}\,] = 0\,, \tag{G.141}$$

with $\alpha, \beta = (1, 2)$, and describing the two angular momentum system by the Jordan-Schwinger maps:

$$\mathbf{J}_1 = \frac{\hbar}{2}\, A^\dagger_1\, \boldsymbol{\sigma}\, A_1\,, \qquad \mathbf{J}_2 = \frac{\hbar}{2}\, A^\dagger_2\, \boldsymbol{\sigma}\, A_2\,. \tag{G.142}$$

Then the total angular momentum operator in occupation number space is given by

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 = \frac{\hbar}{2}\, A^\dagger_1\, \boldsymbol{\sigma}\, A_1 + \frac{\hbar}{2}\, A^\dagger_2\, \boldsymbol{\sigma}\, A_2\,. \tag{G.143}$$

So let us define the tensor product $A^\dagger[(1/2, 1/2)\, S, M]$ by:

$$A^\dagger_{1,2}[(1/2, 1/2)\, S, M] = \sum_{q_1,q_2} \langle\, 1/2, q_1, 1/2, q_2 \,|\, (1/2, 1/2)\, S, M \,\rangle\, A^\dagger_{1\,q_1}\, A^\dagger_{2\,q_2}\,. \tag{G.144}$$

Dropping the 1/2 notation, this is:

$$A^\dagger_{1,2}(S, M) = \begin{cases} \big(\, A^\dagger_{1+}\, A^\dagger_{2-} - A^\dagger_{1-}\, A^\dagger_{2+}\,\big)/\sqrt{2}\,, & \text{for } S = M = 0, \\ A^\dagger_{1+}\, A^\dagger_{2+}\,, & \text{for } S = 1,\ M = +1, \\ \big(\, A^\dagger_{1+}\, A^\dagger_{2-} + A^\dagger_{1-}\, A^\dagger_{2+}\,\big)/\sqrt{2}\,, & \text{for } S = 1,\ M = 0, \\ A^\dagger_{1-}\, A^\dagger_{2-}\,, & \text{for } S = 1,\ M = -1. \end{cases} \tag{G.145}$$

Note that $A^\dagger_{1,2}(0, 0) = \det[\, A^\dagger\,]/\sqrt{2}$. We can also define a mixed tensor $R_{1,2}[(1/2, 1/2)\, S, M]$ by:

$$R_{1,2}[(1/2, 1/2)\, S, M] = \sum_{q_1,q_2} \langle\, 1/2, q_1, 1/2, q_2 \,|\, (1/2, 1/2)\, S, M \,\rangle\, A^\dagger_{1\,q_1}\, A_{2\,q_2}\,. \tag{G.146}$$

Again dropping the 1/2 notation, this is:

$$R_{1,2}(S,M) = \begin{cases} \left(A^\dagger_{1+} A_{2-} - A^\dagger_{1-} A_{2+} \right)/\sqrt{2}\,, & \text{for } S = M = 0, \\ A^\dagger_{1+} A_{2+}\,, & \text{for } S = 1,\ M = +1, \\ \left(A^\dagger_{1+} A_{2-} + A^\dagger_{1-} A_{2+} \right)/\sqrt{2}\,, & \text{for } S = 1,\ M = 0, \\ A^\dagger_{1-} A_{2-}\,, & \text{for } S = 1,\ M = -1. \end{cases} \tag{G.147}$$

We can also define the adjoints of both of these operators, so there are a total of four mixed tensor operators of rank one for the Bose operator representation of angular momentum. In general, these Bose tensor operators are not Hermitian.

Where am I going here and what am I trying to do? Is this section necessary?

References

[1] J. Schwinger, "On angular momentum," (1952). Report NYO-3071, Nuclear Development Associates, Inc., White Planes, New York (unpublished).

ANNOTATION: This paper was never published. However it is reproduced in the collection of articles in Biedenhard and Van Dam[6] which may be the *only* source of this paper.

[2] E. P. Wigner, *Gruppentheorie und ihre Anwendung auf dei Quantenmechanic der Atomspektren* (Braunschweig, Berlin, 1931). English translation: Academic Press, Inc, New York, 1959.

[3] L. C. Biedenharn and J. D. Louck, *Angular momentum in quantum physics: theory and application*, volume 8 of *Encyclopedia of mathematics and its applications* (Addison-Wesley, Reading, MA, 1981).

[4] A. R. Edmonds, *Angular momentum in quantum mechanics* (Princeton University Press, Princeton, NJ, 1996), fourth printing with corrections, second edition.

ANNOTATION: This printing corrected several major errors in Chapter 4 in earlier printings.

[5] R. T. Sharp, "Simple derivation of the Clebsch-Gordan coefficients," Am. J. Phys. **28**, 116 (1960).

[6] L. C. Biedenharn and H. van Dam, *Quantum theory of angular momentum*, Perspectives in physics (Academic Press, New York, NY, 1965).

ANNOTATION: This is a collection of early papers on the theory of angular momentum.

Index