Quantum mechanics and the quark model: An introductory course

Mark W. Paris

Center for Nuclear Studies, The George Washington University
E-mail address: mparis@gwu.edu

Dedicated to the loves of my life, my
daughter, Suzanne Emmy and wife Rachel.

Contents

1. Preface

The purpose of this short textbook is to teach basic elements of quantum mechanics, group theory, and many-body physics to undergraduate and advanced high school students. High school students?! Indeed, the book began as a set of self-guided study notes to teach advanced high school students in an honors research/mentorship program in southeastern Virginia how to calculate the mass splitting of the nucleon, $N(939)$ and lightest spin-$\frac{3}{2}$, isospin-$\frac{3}{2}$ state, the $\Delta(1232)$ due to the hyperfine interaction between the quarks. (If you don't know what this means, don't worry, after you read and work through the problems in these 60-some odd pages, you will.) The students who have worked through the book were, I think, successful in attaining that objective.

The book – honestly, calling it a book is a bit of a stretch in its current form, but it's a little more than a set of notes because it attempts to teach the reader as the exercises are worked out – presupposes only a knowledge in complex arithmatic and calculus sufficient to integrate functions of single variable and understand very basic linear differential equations. The style of the book is that of a self-study guide. I have gone to some, though no doubt insufficient, length at making the technical development of the narrative as accessible as possible while still requiring that the significant logical developments be left to the reader to work out. In other words, there are a lot of exercises that are not difficult and occasionally there's a tough problem that brings together the results of the exercises that came before. I've also tried to include much of the algebra and intermediate steps to working out the exercises, either immediately before or just after the given exercise, directly in the text. No doubt more could have been included, but there is a fine line between showing not enough and too much. I'm sure I have crossed it numerous times and would appreciate any comments on the text, of a general nature, to `mparis@gwu.edu`.

My gratitude goes to the two students, Kiralyse Gonzalez and Joel Gillespie, without whom this work would likely not have happened. As they worked through the material, they provided valuable feedback and constructive analysis of the text itself, often pointing out difficult sections and mistakes. The text itself is far from a complete work, but I hope it will be useful for people who want to learn some quantum mechanics in a 'quick-start' format. Of course, as is usual with the quick-start approach, it can only offer an entrée into a deeper and more comprehensive knowledge of the material. You won't, in short, be an expert in quantum physics after reading this book. You will, however, have learned some valuable calculational tools and analytical techniques that you will encounter frequently in the study of quantum systems.

Mark Paris
Washington, D.C.
2010 September 11

> *Get over the idea that only children should spend their time in study. Be a student so long as you still have something to learn, and this will mean all your life. –Henry L. Doherty*

CHAPTER 1

Group theory

The ultimate objective of this book, as mentioned in the *Preface* is a calculation of the mass splitting of the nucleon $P_{11}(939)$ and the delta-baryon $\Delta_{33}(1232)$ due to the hyperfine interaction. (Don't worry about what those symbols mean for now. We'll get to all that in short order.) In order to achieve this ultimate objective, we'll need to learn some mathematical tools as well as some physics.

The first mathematical tool we'd like to take into our hands is that of group theory. Group theory is simply the mathematical expression of the fact that, roughly speaking, if you do certain operations twice, it's the same thing as doing the operation once, but perhaps in a slightly modified fashion. Think of moving (translating) an object from point A to point B. And then from B to C. It's the same as just translating it from A to C directly. This idea works not just for translations, but for rotations and other transformations, as well.

We now turn to the rather abstract context of rotations of quantum objects like electrons. Of course, we'll relate this to more familiar contexts as we go along.

1. Spinors

In a very loose and not-rigorous sense, group theory is the generalization of the algebra of real numbers to an algebra of matrices. (I should caution you though that group theory is more abstract than this and doesn't really depend on matrices at all.)

Suppose we have a pair of complex numbers ψ_1, ψ_2 which we write as a sort of vector:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}. \tag{1.1}$$

Further suppose that all we're concerned with, for now, is the **norm**[1] (think of it as a sort of generalized length) of this vector. The norm $|\psi|^2$ is defined by

$$|\psi|^2 = \psi_1^* \psi_1 + \psi_2^* \psi_2, \tag{1.2}$$

where ψ_1^* is the **complex conjugate** of $\psi_1 = \psi_{1,r} + i\psi_{1,i}$ defined by

$$\psi_1^* = \psi_{1,r} - i\psi_{1,i}, \tag{1.3}$$

[1] New math terms or definitions appear bold.

where $\psi_{1,r}, \psi_{1,i}$ are the real and imaginary parts of the *complex* number ψ_1.[2] We have a short-hand way of writing the expression for the norm:

$$\psi^\dagger = (\psi_1^*, \psi_2^*) \tag{1.4}$$

$$|\psi|^2 = \psi^\dagger \psi \tag{1.5}$$

$$= (\psi_1^*, \psi_2^*) \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

$$= \psi_1^* \psi_1 + \psi_2^* \psi_2$$

$$= |\psi_1|^2 + |\psi_2|^2 . \tag{1.6}$$

The superscripted "dagger," † instructs us to take the complex conjugate of each **element** or **component**, ψ_1 or ψ_2 in the vector (or, in case we're dealing with matrices, each element of the matrix) and then **transpose** it. The "transpose" operation means 'interchange the rows and the columns.' The composite operation ("complex conjugate & transpose" or "transpose & complex conjugate", the order doesn't matter) is called **Hermitian conjugation**.

Exercise 1.1. Starting with the complex vector

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \tag{1.7}$$

show that we end up with the matrix

$$\psi^\dagger = (\psi_1^*, \psi_2^*) , \tag{1.8}$$

whether we do:

(i) complex conjugate *then* transpose;
(ii) transpose *then* complex conjugate.

This is how we calculate the norm of a two-dimensional complex vector. (Actually, you can generalize everything we've covered so far to the case of N-dimensional complex vectors.) Two-dimensional complex vectors have a name – **spinors**. The norm of a spinor is an important physical quantity – I won't tell you why just now, but just remember that we're learning how to calculate it because it's important for the physics of quantum objects like electrons.

Now, since only the norm of ψ is important, we can 'play around' with the components ψ_1, ψ_2 in a certain way without changing the value of the norm. This 'certain way' is, more specifically, taking linear combinations of the ψ_1 and ψ_2. For example, $-5\psi_1 + (8 + i\pi)\psi_2$ is one such linear combination of the spinor's components among infinitely many such linear combinations – we could have chosen any complex numbers in place of -5 or $(8 + i\pi)$. Since the original spinor, ψ has two components, ψ_1 and ψ_2, we need to form two such linear combinations so that the new spinor which we've 'played' with, or **transformed**, is still a two-dimensional complex vector. We can codify this mathematically using a "2-by-2" complex matrix. Here's how.

Suppose I have a **square matrix** U with two rows and two columns (a "2-by-2" matrix) with complex numbers a, b, c, d for entries:

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} . \tag{1.9}$$

[2] I'm assuming that you know how to do complex arithmetic and you know that $i \equiv \sqrt{-1}$.

(It's "square" because the number of the number or rows and columns are the same.) Now, take this matrix and multiply it from the left (with matrix multiplication) by the spinor ψ:

$$\begin{aligned} U\psi &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \\ &= \begin{pmatrix} a\psi_1 + b\psi_2 \\ c\psi_1 + d\psi_2 \end{pmatrix} = \psi', \end{aligned} \tag{1.10}$$

where we've given this new spinor a name, ψ'. So you see that we've formed *two* linear combinations, one for the upper component $(a\psi_1 + b\psi_2)$ and one for the lower component $(c\psi_1 + d\psi_2)$, by multiplying the spinor ψ by a single "2-by-2" matrix.

(I'm assuming that you know how to multiply matrices here. By way of a very brief review, if the matrix C is a product of two matrices A and B, $C = AB$ then in terms of components is

$$C_{ik} = \sum_{j=1}^{N} A_{ij} B_{jk} \tag{1.11}$$

where C is an $C_r \times C_c$ matrix, A is a $C_r \times N$ matrix, B is a $N \times C_c$ matrix. A spinor is a 2-by-1 matrix. Also remember that matrix multiplication is not **commutative** – this means that for matrices $AB \neq BA$.)

Let's work this out first without specific reference to matrices. We're studying the transformation of the norm under the action of the matrix U so we now calculate the norm of the new, transformed spinor:

$$\begin{aligned} |\psi'|^2 &= |\psi_1'|^2 + |\psi_2'|^2 \\ &= (a\psi_1 + b\psi_2)^*(a\psi_1 + b\psi_2) + (c\psi_1 + d\psi_2)^*(c\psi_1 + d\psi_2) \\ &= (a^*\psi_1^* + b^*\psi_2^*)(a\psi_1 + b\psi_2) + (c^*\psi_1^* + d^*\psi_2^*)(c\psi_1 + d\psi_2) \\ &= (|a|^2 + |c|^2)\,|\psi_1|^2 + (|b|^2 + |d|^2)\,|\psi_2|^2 \\ &+ (a^*b + c^*d)\psi_1^*\psi_2 + (ab^* + cd^*)\psi_1\psi_2^*. \end{aligned} \tag{1.12}$$

Clearly, for arbitrary complex numbers a, b, c & d the norm of the transformed spinor $|\psi'|^2$ is not the same as the norm of the original spinor $|\psi|^2$, since the last two terms of the last line of Eq.(1.12) don't appear in the expression for $|\psi|^2$. But what if we *require* that $|\psi|^2 = |\psi'|^2$. This places **constraints** among a, b, c & d. The next exercise asks you to produce the equations expressing these constraints:

$$|a|^2 + |c|^2 = 1 \tag{1.13a}$$

$$|b|^2 + |d|^2 = 1 \tag{1.13b}$$

$$a^*b + c^*d = 0 \tag{1.13c}$$

$$ab^* + cd^* = 0. \tag{1.13d}$$

(The last equation is just the complex conjugate of the second-to-last one, and so doesn't have any independent information.)

Exercise 1.2. Derive Eqs.(1.13) by writing down Eq.(1.12) (doing the algebra for yourself) and requiring $|\psi|^2 = |\psi'|^2$.

Using these equations we can derive the relations

$$|b|^2 = |c|^2 \tag{1.14}$$

$$|a|^2 = |d|^2 . \tag{1.15}$$

EXERCISE 1.3. Derive these from the equations above.

This is a good time to introduce a more compact and advanced notation. It will pay us later to invest a little time now learning short-hand ways of writing the expression for the norm. Consider the norm of the transformed spinor, $|\psi'|^2$. We'll start with the definition of the norm from Eq.(1.6) and substitute in the transformed spinor in terms of the transformation matrix U and the original spinor, ψ:

$$\begin{aligned} |\psi'|^2 &= (\psi')^\dagger \psi' \\ &= (U\psi)^\dagger U\psi \qquad (1.16) \\ &= \psi^\dagger U^\dagger U\psi. \qquad (1.17) \end{aligned}$$

Let's examine the last line of the above derivation. To get there we used

$$(U\psi)^\dagger = \psi^\dagger U^\dagger. \tag{1.18}$$

EXERCISE 1.4. Prove Eq.(1.18) by first calculating $(U\psi)^\dagger$ and then $\psi^\dagger U^\dagger$ with U and ψ given by

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad\qquad \psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \tag{1.19}$$

showing that they are equal.

Continuing from Eq.(1.17) and using the fact that the transformation *induced* on ψ by U doesn't change the norm, we get a constraint on the matrix U:

$$|\psi'|^2 = |\psi|^2 \tag{1.20}$$

$$\implies \psi^\dagger U^\dagger U\psi = \psi^\dagger \psi \tag{1.21}$$

$$\implies U^\dagger U = \mathbb{1}_2. \tag{1.22}$$

Here, $\mathbb{1}_2$ is the unit matrix in two dimensions (denoted by the subscript '2'), which has the explicit form

$$\mathbb{1}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.23}$$

The important point here is that Eq.(1.22) contains the same information as the constraints expressed in Eqs.(1.13) which were obtained by working out $|\psi'|^2$ as in Eq.(1.12). Equation (1.22) is, incidentally, referred to as a **unitarity constraint** on the transformation matrix, U and defines U as a **unitary matrix**. Let's see in detail how the unitarity constraint on U had the same information as Eqs.(1.13) by working out the following exercise.

EXERCISE 1.5. Show that the matrix equation $U^\dagger U = \mathbb{1}_2$, written explicitly as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^\dagger \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{1.24}$$

gives Eqs.(1.13) by: *i)* performing the Hermitian conjugate action (†) on the left-most matrix above; *ii)* multiplying the resulting matrices; *iii)* equating each of the four elements of the left-and right-hand sides of the equality.

The obvious advantage in this more compact matrix notation is that we never had to 'open-up' the object U to get the constraint Eq.(1.22). Also, we never had to refer to the components of ψ, ψ_1 and ψ_2. The convenience of the compact matrix notation is obvious. Note, however, that we only introduced the matrix notation *after* we got some experience working with the component notation. This is a good rule to follow generally – when tempted to work with compact notation, we should make sure that we understand how it works in terms of the quantities and elements that comprise the short-hand forms. This way we're sure not to misinterpret their meanings and make a mistake in their manipulation.

Enough pedagogical advice – let's use what we've learned so far to simplify the form of the transformation matrix, U. Doing this will allow us to more simply understand the transformation in terms of group theory and to see why group theory forces U to take this specific simplified form. What we mean by 'simplify the form of U' is this – we'll rewrite U in terms of the fewest number of its complex number elements a, b, c, and d as possible. Since we have the relationships between these elements provided by the constraints of Eqs.(1.13), we can eliminate some of these elements in favor of others. Actually, it's a little easier to do this simplifying algebra if we introduce another quantity – the **determinant** of U.

The determinant of a 2-by-2 matrix M is defined by

$$M = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix} \tag{1.25}$$

$$\det M = m_1 m_4 - m_2 m_3. \tag{1.26}$$

The determinant, which is likely familiar to you already, is so important that we'll pause for a digression on its properties. In the following subsection, we prove some important properties of the determinant before returning to the issue of simplifying U in terms of its elements.

1.1. Properties of determinants. Recall the equation for the determinant of a 2-by-2 matrix:

$$M = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix} \tag{1.27}$$

$$\det M = m_1 m_4 - m_2 m_3. \tag{1.28}$$

The determinant is just a way of getting a *single, complex number* from a square matrix. It has the following properties:

a) $\det U_1 U_2 = \det U_1 \det U_2$;
b) $\det U^T = \det U$;
c) $\det U^\dagger = (\det U)^*$;
d) $\det aU = a^2 \det U$ (for *rank 2* **only**!)
e) $\det U^{-1} = \frac{1}{\det U}$;
f) $\det \begin{pmatrix} a & b \\ kc & kd \end{pmatrix} = k \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ (common row factor);
g) $\det \begin{pmatrix} ka & b \\ kc & d \end{pmatrix} = k \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ (common column factor).

Please note that all of these (except the one noted exception) are applicable for rank-n (n-by-n) matrices.

EXERCISE 1.6. Prove the above properties of the determinant. **Extra credit**: generalize $\det aU$ to the case where U is a *rank–n* matrix.

Resuming our pursuit of the simplest form for U allowed by the unitarity constraint on U, Eq.(1.22) (or equivalently Eqs.(1.13)), we'll first need to use a property of the determinant that we just proved.

EXERCISE 1.7. Using property 'c' of the determinant and Eq.(1.22), show that

$$|\det U|^2 = 1. \tag{1.29}$$

EXERCISE 1.8. Show that the determinant of U in Eq.(1.31) can be expressed in either of the two forms

$$\det U = -\frac{b}{c^*}, \qquad \det U = \frac{a}{d^*}, \tag{1.30}$$

using Eqs.(1.13a–1.13c).

Using these we can rewrite **transformation matrix** U as

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}. \tag{1.31}$$

EXERCISE 1.9. Using this form of the matrix U, multiply it (from the left, of course) on ψ, use Eqs.(1.13) and show that the norm of the resulting spinor is the same as $|\psi|^2$.

Before we move on to the next section and consider group theory in a more general sense, let's have a look at what we've done. We started with the matrix U in Eq.(1.9) which has 4 complex numbers, each of which has two independent real components for a total of 8 independent real numbers making up the matrix U. Then we require that the matrix U doesn't change ("leaves invariant") the norm of the spinor ψ, $|\psi|^2$. This imposes four conditions on the elements of U: Eqs.(1.13a,1.13b) are each one condition since they are relations among real – not complex – quantities, while Eq.(1.13c) gives two conditions since it's a relation among complex quantities. So we're down to just 4 independent real numbers – these are just the real and imaginary parts of a and b the only complex numbers that appear in Eq.(1.31) for U. (Note that complex numbers are not independent of their complex conjugates, eg. a^* isn't independent of a.) The fact that the determinant of U, $\det U = 1$ (as you proved in Ex.(1.8)) is due to the fact that $|a|^2 + |b|^2 = 1$. So you can see that a and b have one relation between them and therefore there are only 3 independent real numbers in U. (I have cheated here – Eqs.(1.13) fix the ratios of b/c^* and a^*/d only up to a **phase factor**, $e^{i\alpha}$, and I've arbitrarily chosen $\alpha = 0$. We'll return to this point at the end of the chapter.

2. Group properties

Now that we've some some "real" world (that is, real "*quantum mechanical world*") examples of groups ($U(2)$ and $SU(2)$) and the objects they operate on (*spinors*) let's firm up the mathematical definitions of these **sets**. The essential properties of the abstract group consisting of a **set** of elements and given a **group product** are:

- **closure**: the product of two elements is itself in the set;
- **associative**: $U_1(U_2U_3) = (U_1U_2)U_3$;

- **unit element 𝟙**: $\mathbb{1}U = U\mathbb{1} = U$;
- **inverse** U^{-1}: $UU^{-1} = U^{-1}U = \mathbb{1}$.

We can show that the objects $U(2)$ and $SU(2)$ are, in fact, groups by considering representatives from each of these groups and checking that they satisfy the specific properties above. The only property that we didn't explicitly check was the associtative rule. But it's a fact that matrix multiplication satisfies the associative rule and we won't explicitly check this.

Let's have a quick review of $U(2)$ and $SU(2)$.

2.1. $U(2)$**.** A general form for the elements of $U(2)$ is

$$U = \begin{pmatrix} a & b \\ -b^* e^{i\alpha} & a^* e^{i\alpha} \end{pmatrix} \tag{1.32}$$

$$\det U = e^{i\alpha} \tag{1.33}$$

and we have proved that these elements satisfy the **unitarity** condition

$$U^\dagger = U^{-1}. \tag{1.34}$$

Recall that these equations define elements of the group $U(2)$ of "unitary matrices of rank 2." Let's check that U's of the form of Eq.(1.32) (that is, $U \in U(2)$ – this reads: 'U is an **element** of $U(2)$') are **closed** (see the **closure** property above) under matrix multiplication.

Exercise 1.10. Suppose $U_1, U_2 \in U(2)$.

$$U_1 = \begin{pmatrix} a_1 & b_1 \\ -b_1^* e^{i\alpha_1} & a_1^* e^{i\alpha_1} \end{pmatrix} \qquad U_2 = \begin{pmatrix} a_2 & b_2 \\ -b_2^* e^{i\alpha_2} & a_2^* e^{i\alpha_2} \end{pmatrix}. \tag{1.35}$$

Show that the resulting matrix, $U = U_1 U_2$ satisfies $U \in U(2)$. In order to do this, you need to show that: *i*) $\det U = e^{i(\alpha_1+\alpha_2)}$ and *ii*) Eqs.(1.13) are satisfied for the elements of U.

2.2. $SU(2)$**.** The general form for the elements of $SU(2)$ are obtained from that of $U(2)$ by setting $\alpha = 0 \implies e^{i\alpha} = 1$:

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \tag{1.36}$$

$$\det U = 1. \tag{1.37}$$

These equations define the elements of $SU(2)$ the group of "special unitary matrices of rank 2."

Exercise 1.11. Repeat Ex.(1.10) for $SU(2)$ using the results from Ex.(1.10). Don't multiply matrices out! Be clever.

3. Lie algebra

Now, let's specialize to the group $SU(2)$. This is what's called a **Lie group** (named after the Norwegian mathematician Sophus Lie who lived in the latter half of the nineteenth century). We've shown that the form of the elements of $SU(2)$ can be written

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \tag{1.38}$$

and that the condition $\det U = 1$ means that there are three independent real numbers or **parameters** within the complex numbers a and b. Our goal in this section is to use the properties of the **Lie algebra**, denoted $\mathfrak{su}(2)$, to write the matrix U of Eq.(1.38) in the form

$$U = e^{i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}}. \tag{1.39}$$

This is probably something new to you – we're taking the exponential function of an argument which isn't a real number, as we're used to doing, but a *matrix*, $\boldsymbol{\sigma}\cdot\boldsymbol{\theta}$. (Don't worry about exactly what $\boldsymbol{\sigma}\cdot\boldsymbol{\theta}$ is right now. It'll become clear as we go on. Just be aware that it's a *rank–2* matrix.)

Let's rewrite Eq.(1.38), separating the real and imaginary parts of a and b. We write:

$$U = a_r \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + ib_i \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + ib_r \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + ia_i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.40}$$

$$= a_r \mathbb{1} + i\mathbf{u}\cdot\boldsymbol{\sigma} \tag{1.41}$$

where we've defined

$$\mathbf{u} = \begin{pmatrix} b_i \\ b_r \\ a_i \end{pmatrix} \qquad \boldsymbol{\sigma} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} \tag{1.42}$$

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{1.43}$$

where we've used the definition for the **scalar product** (or 'dot product')

$$\mathbf{u}\cdot\boldsymbol{\sigma} = u_1\sigma_1 + u_2\sigma_2 + u_3\sigma_3. \tag{1.44}$$

Exercise 1.12. Work out Eqs.(1.40,1.41).

A lot has happened here, so let's sift through the pile carefully. The first thing we did was to break up the matrix U into real and imaginary components of the parameters a, b. You see that you get the unit matrix $\mathbb{1}$, and after factoring out an $i = \sqrt{-1}$ we get three more matrices $\sigma_1, \sigma_2, \sigma_3$ which are often written as a vector, $\boldsymbol{\sigma}$. (In fact sometimes the labels 1,2,3 are written as x, y, z.) These are a special set of matrices that arise *all the time* in the physics of spin-$\frac{1}{2}$ particles. They're called the **Pauli matrices** and they have special properties. Let's write these down and then I'll explain them:

$$\begin{aligned} [\sigma_1, \sigma_2] &= 2i\sigma_3 \\ [\sigma_2, \sigma_3] &= 2i\sigma_1 \\ [\sigma_3, \sigma_1] &= 2i\sigma_2. \end{aligned} \tag{1.45}$$

This is the **Lie algebra**. We have a new object here: [,] called the **commutator**. It's defined as

$$[A, B] = AB - BA \tag{1.46}$$

and has the property that

$$[A, B] = -[B, A]. \tag{1.47}$$

Clearly, if either (or both) A and B are real numbers the commutator is zero. Since matrix multiplication isn't commutative, the commutator of matrices is, in general, not zero. (Though there are other objects which are not matrices that

don't commute. Just a warning that commutators are useful for things other than matrices, but we'll confine our attention to just matrices for now.)

EXERCISE 1.13. Prove Eq.(1.47) from the definition in Eq.(1.46).

EXERCISE 1.14. Prove Eqs.(1.45).

So you can see from Eqs.(1.45) that there is a sort of closure of the Pauli matrices under the action of the commutator. I'm not saying that the Pauli matrices form a group. They don't. They are the elements of a special sort of vector space called a Lie algebra. Just keep this in the back of your mind for now.

EXERCISE 1.15. Prove that the Pauli matrices $\{\sigma_1, \sigma_2, \sigma_3\}$ don't form a group.

In fact, going back to Eq.(1.40), what we have done is write the general element of $SU(2)$ in terms of a **basis** of matrices consisting of the identity matrix and the Pauli matrices. And any element in $SU(2)$ can be written this way in terms of the parameters a_r, b_i, b_r, a_i and the basis matrices $\{\mathbf{1}, \boldsymbol{\sigma}\}$. (Remember that there are only three independent paramters since they're related by

$$a_r^2 + b_i^2 + b_r^2 + a_i^2 = 1. \tag{1.48}$$

Notice something else about the Pauli matrices. They're **traceless**. If we have an n-by-n matrix A then the **trace** is defined as

$$\text{Tr } A = \sum_{i=1}^{n} A_{ii}, \tag{1.49}$$

just the sum of the diagonal elements. So the Pauli matrices have

$$\text{Tr } \sigma_i = 0. \tag{1.50}$$

EXERCISE 1.16. Using Eq.(1.36) calculate the Tr U.

EXERCISE 1.17. Using Eq.(1.41) calculate the Tr U. Compare to your result from Ex.(1.16). Do they agree?

Now, onto the *coup de grace* – proving Eq.(1.39). Suppose that we write

$$a_r = \cos\theta \tag{1.51}$$

then Eq.(1.48) is satisfied if we take

$$\mathbf{u} = \hat{\boldsymbol{\theta}} \sin\theta \qquad \text{(recall Eq.(1.42))} \tag{1.52}$$

$$\hat{\boldsymbol{\theta}} = \frac{\boldsymbol{\theta}}{\theta} \tag{1.53}$$

$$\boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} \tag{1.54}$$

$$\theta = \sqrt{\theta_1^2 + \theta_2^2 + \theta_3^2}. \tag{1.55}$$

Then we have, if we substitute into Eq.(1.41)

$$U = \cos\theta + i\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\theta}} \sin\theta \tag{1.56}$$

and now, we've renamed our independent parameters from a_r, b_r, b_i to $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ (or $\boldsymbol{\theta} = (\theta_x, \theta_y, \theta_z)$). The symbol $\hat{\boldsymbol{\theta}}$ is the symbol for a unit vector which points in the direction of the axis about which we'll rotate the system:

$$\boldsymbol{\theta} = (\theta_x, \theta_y, \theta_z) \tag{1.57}$$
$$= \theta_x \hat{\mathbf{x}} + \theta_y \hat{\mathbf{y}} + \theta_z \hat{\mathbf{z}} \tag{1.58}$$
$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}/\theta \tag{1.59}$$
$$= \hat{\theta}_x \hat{\mathbf{x}} + \hat{\theta}_y \hat{\mathbf{y}} + \hat{\theta}_z \hat{\mathbf{z}} \tag{1.60}$$
$$\hat{\theta}_x = \theta_x/\theta \qquad \hat{\theta}_y = \theta_y/\theta \qquad \hat{\theta}_z = \theta_x/\theta \tag{1.61}$$
$$\theta = \sqrt{\theta_x^2 + \theta_y^2 + \theta_z^2}. \tag{1.62}$$

Let's work through this with some problems.

EXERCISE 1.18. Determine the variables (b_i, b_r, a_i) in terms of $(\theta, \hat{\theta}_x, \hat{\theta}_y, \hat{\theta}_z)$.

EXERCISE 1.19. Prove that Eq.(1.48) is satisfied by Eqs.(1.51,1.52).

EXERCISE 1.20. Write out Eq.(1.56) in terms of its elements. Each element is a cos or sin function muliplied by coefficients which are complex numbers formed from the components of $\hat{\boldsymbol{\theta}}$.

EXERCISE 1.21. When we wrote $a_r = \cos\theta$ we assumed that $|a_r| \leq 1$ (why do I say this?). Prove that this indeed the case from the condition $\det U = 1$.

EXERCISE 1.22. Calculate the determinant of U using Eq.(1.56).

EXERCISE 1.23. Calculate the trace of U using Eq.(1.56). Compare to the results for Ex.(1.16) and Ex.(1.17).

Now we turn to the formidable task of proving Eq.(1.39). We know how to take exponential functions of real arguments. Simply raise the number $e = 2.7813$ to that power. Rigorously though, any function is defined by its series expansion. Given a function whose derivatives all exist in an **open neighborhood** of x, Taylor's theorem says that

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n, \tag{1.63}$$

where $f^{(n)}(x) = \frac{d^n f(x)}{dx^n}$ and $f^{(n)}(0) = f^{(n)}(x)\big|_{x=0}$.

EXERCISE 1.24. Using Taylor's theorem, prove that the series expansion of $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.

Now that we know, rigorously, how to take the exponential function of a real number, x, let's think about how to take the exponential function of a complex number, z. Letting $z = x + iy$ we have

$$e^z = e^x e^{iy}. \tag{1.64}$$

We already have done e^x. What about e^{iy}? This is a weird thing. if I write it out numerically for, say $y = 3$ then we have

$$e^{iy} = (2.7813)^{3\sqrt{-1}} \tag{1.65}$$

and it's not at all clear what to do with this. We do know, however, how to multiply (and divide, etc.) complex numbers. That suggests we might get somewhere if we use the series representation for the exponential function from Ex.(1.24).

EXERCISE 1.25. From $e^{iy} = \sum_{n=0}^{\infty} \frac{(iy)^n}{n!}$ prove that $e^{i\alpha} = \cos\alpha + i\sin\alpha$.

Now that we know how to exponentiate complex numbers like $z = x + iy$, we can ask: how do we exponentiate *matrices* as in Eq.(1.39)? Answer: go to the series representation!

First we need some background on how to take products (ie., multiply) the Pauli matrices.

EXERCISE 1.26. Prove that $\sigma_i\sigma_i = 1$ for $i = 1, 2, 3$.

This next exercise is simply a more concise restatement of Eqs.(1.43).

EXERCISE 1.27. Show that $[\sigma_i, \sigma_j] = 2i\sum_{k=1}^{3} \varepsilon_{ijk}\sigma_k$. Here $\varepsilon_{ijk} = 1$ if $ijk = 123, 231, 312$, $\varepsilon_{ijk} = -1$ if $ijk = 132, 321, 213$, and $\varepsilon_{ijk} = 0$ if $i = j$ or $j = k$ or $k = i$. The ε_{ijk} is called the Levi-Civita tensor.

EXERCISE 1.28. Using Exs.(1.26) and (1.27) show that $\sigma_i\sigma_j = \delta_{ij} + i\sum_{k=1}^{3} \varepsilon_{ijk}\sigma_k$. Here $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.

EXERCISE 1.29. Now, using the last four problems we can show that $U = e^{i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}}$.

4. Rotations in 2 and 3 dimensions

This section is a bit of a detour, a *denouement* from the main track of the chapter. It serves as a conduit to understanding what spinors are and what the transformation matrix of Eq.(1.39) (or Eq.(1.56)) does to spinors. It does this by looking at rotations of good old regular vectors in two and three dimensions.

What we've done so far is to figure out the structure of the elements of $U(2)$, $SU(2)$ and we briefly had a look at the Lie algebra of $SU(2)$ (which is often denoted $\mathfrak{su}(2)$). But what are these things? Well, just confining our attention to $SU(2)$, we'll find out soon enough that these elements (like $U(\boldsymbol{\theta})$ of Eq.(1.56)) correspond to rotations of things described by Pauli spinors, like electrons and quarks and nucleons, and Σ^- hyperons – any spin$-\frac{1}{2}$ particles, in fact.

But first, let's study how rotations work in two and three dimensions; that is, how do vectors in two and three dimensions **transform** under rotation?

Starting with two dimensions (2D), we have a vector

$$\mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{1.66}$$

$$x = |\mathbf{r}| \cos\theta, \quad y = |\mathbf{r}| \sin\theta \tag{1.67}$$

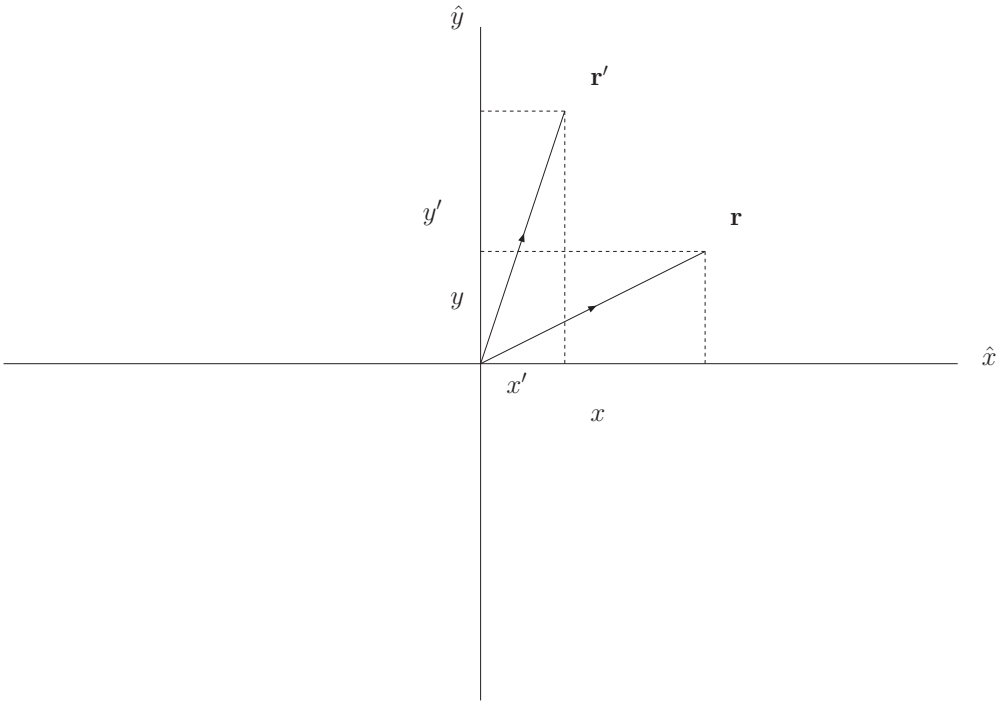

FIGURE 1. The rotation of the vector $\mathbf{r}$ to $\mathbf{r}'$.

which is rotated to

$$\mathbf{r}' = \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{1.68}$$

$$x' = |\mathbf{r}| \cos\theta', \quad y' = |\mathbf{r}| \sin\theta' \tag{1.69}$$

where the angles θ and θ' are measured with respect to the $\hat{\mathbf{x}}$ axis. See Fig.(1). Note that the same length, $|\mathbf{r}|$ appears for both vectors $\mathbf{r}$ and $\mathbf{r}'$ since rotations don't change the lengths of vectors. They merely turn them while keeping their length fixed.

EXERCISE 1.30. Prove Eqs.(1.69) by looking at Fig.(1) and assuming $|\mathbf{r}'| = |\mathbf{r}|$. The angle between $\mathbf{r}'$ and the $x-$axis is θ'.

We can get from $\mathbf{r}$ to $\mathbf{r}'$ ("**transform** $\mathbf{r}$ to $\mathbf{r}'$") by applying a matrix, an element of $SO(2)$ in fact (we'll find out what the 'O' means in a minute) called $R(\phi)$:

$$\mathbf{r}' = R(\phi)\mathbf{r} \tag{1.70}$$

$$R(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \tag{1.71}$$

where we can use popular trigonometric identities (angle addition formulas) to prove that

$$\phi = \theta' - \theta. \tag{1.72}$$

So ϕ is angle *between* $\mathbf{r}$ and $\mathbf{r}'$.

EXERCISE 1.31. Prove Eq.(1.72).

Now that we have the matrix $R(\phi)$ we can have a look at some of its properties.

EXERCISE 1.32. Show that $R(\phi)$ has $\det R = 1$.

EXERCISE 1.33. Show that $R^{-1}(\phi) = R^T(\phi) = R(-\phi)$ by applying the general formula for the inverse of a 2-by-2 matrix that you've derived earlier. Note that there are two equalities to prove.

The problem in Ex.(1.33) shows that the inverse of the matrix $R(\phi)$ is equal to its transpose. (Remember that transpose just means: interchange rows and columns.) And this is where the 'O' comes from in $SO(2)$ meaning "orthogonal."[3] That is: **orthogonal** $\Leftrightarrow R^{-1} = R^T$. So $SO(2)$ is the group of matrices of 'special orthogonal' matrices of rank 2.

EXERCISE 1.34. Where does the 'S' meaning 'special' come from? Show that $|\mathbf{r}'|^2 = \mathbf{r}' \cdot \mathbf{r}'$ is equal to $|\mathbf{r}|^2$ because of this.

EXERCISE 1.35. Prove that these matrices form a group.

EXERCISE 1.36. Show that this group is Abelian, meaning $R_1 R_2 = R_2 R_1$ by considering

$$R_1 = \begin{pmatrix} \cos\phi_1 & -\sin\phi_1 \\ \sin\phi_1 & \cos\phi_1 \end{pmatrix}, \qquad R_2 = \begin{pmatrix} \cos\phi_2 & -\sin\phi_2 \\ \sin\phi_2 & \cos\phi_2 \end{pmatrix}.$$

[3]Strictly speaking, orthogonal transformations preserve the inner product.

That $SO(2)$ is Abelian means that the operation of rotations about a single axis (the rotations we have been considering here are about the axis perpendicular to the $x-y$ plane) **commute** with each other – the order in which we do two rotations doesn't matter. We end up with the same vector (starting with the same vector initially) whatever order we choose to do the rotations in. This will *not* be the case in three dimensions where rotations *do not commute*.

Now that we have the rotation matrices, $R(\phi)$ in two dimensions, it makes the difficult task of understanding the three dimensional (3D) analogs a little easier. Again we start with the 3D vector $\mathbf{r}$

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{1.73}$$

$$x = |\mathbf{r}| \sin\theta \cos\phi, \quad y = |\mathbf{r}| \sin\theta \sin\phi, \quad z = |\mathbf{r}| \cos\theta \tag{1.74}$$

which is rotated to

$$\mathbf{r}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \tag{1.75}$$

$$x' = |\mathbf{r}| \sin\theta' \cos\phi', \quad y' = |\mathbf{r}| \sin\theta' \sin\phi', \quad z' = |\mathbf{r}| \cos\theta' \tag{1.76}$$

where the angles θ and ϕ are measured with respect to the $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ axes, respectively. See Fig.(2), which shows only the vector $\mathbf{r}$. The rotation still takes $\mathbf{r} \to \mathbf{r}'$ but the figure is too messy if we put both vectors on it. (And I couldn't find a picture of it on the web!)

The 3D rotation matrix, which we will call $R(\boldsymbol{\theta})$, is more complicated than the 2D case. In 2D there is only one rotation axis (perpendicular to the $x-y$ plane) and the group structure is particularly simple. In 3D, there are an infinite number of possible directions or axes about which we may rotate a given vector. This makes writing down a single, simple closed form for the rotation matrix, $R(\boldsymbol{\theta})$ more difficult. I'll just show it so you can say you've seen it. We now have to specify

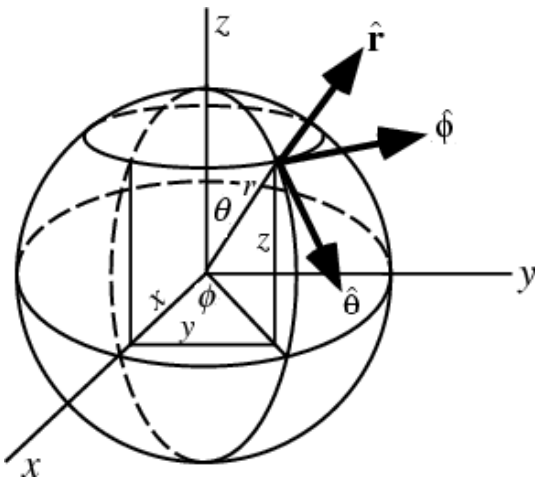

FIGURE 2. The vector $\mathbf{r}$ in spherical coordinates.

a vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) = (\theta_x, \theta_y, \theta_z)$ (we use $(1,2,3)$ or (x,y,z) interchangeably) since we're specifying a rotation about a particular *axis*:

$$\mathbf{r}' = R(\boldsymbol{\theta})\mathbf{r} \tag{1.77}$$

$$R(\boldsymbol{\theta}) = e^{i\mathbf{J}\cdot\boldsymbol{\theta}} \tag{1.78}$$

$$\mathbf{J}\cdot\boldsymbol{\theta} = J_x\theta_x + J_y\theta_y + J_z\theta_z \tag{1.79}$$

$$J_x = i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_y = i\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad J_z = i\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{1.80}$$

The $R(\boldsymbol{\theta})$ is an exact analog to the $U(\boldsymbol{\theta})$ [Eqn.(1.56)] of $SU(2)$. (Except here, we are rotating the a *vector in 3D space* and in the previous sections we were rotating a *spinor in spinor space*.)

The components of the vector $\mathbf{J}$ satisfy the Lie algebra

$$[J_i, J_j] = i\varepsilon_{ijk}J_k \tag{1.81}$$

and the rotation matrix is orthogonal

$$R^{-1} = R^T. \tag{1.82}$$

EXERCISE 1.37. Show that $[J_x, J_y] = iJ_z$.

EXERCISE 1.38. Show that $R^{-1}(\boldsymbol{\theta}) = R^T(\boldsymbol{\theta}) = R(-\boldsymbol{\theta})$. Don't write out the J_i's. You can use the fact that $(e^A)^T = e^{A^T}$ to do this simply. It should only take you three or four lines.

Rotations about the axes x or y or z through an angle θ are simple. You can derive the expressions for these rotations in the same way you showed that Eq.(1.56) can be written

$$e^{i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}} = \cos\theta + i\boldsymbol{\sigma}\cdot\hat{\boldsymbol{\theta}}\sin\theta. \tag{1.83}$$

EXERCISE 1.39. Show that:

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \quad R_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix},$$

$$R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & -\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.84}$$

by taking $\boldsymbol{\theta} = \theta\hat{\mathbf{x}}$ for R_x, $\boldsymbol{\theta} = \theta\hat{\mathbf{y}}$ for R_y, and $\boldsymbol{\theta} = \theta\hat{\mathbf{z}}$ for R_z in Eq.(1.78).

Let's show explicitly that rotations in 3D don't commute.

EXERCISE 1.40. Using the angle $\theta = \frac{\pi}{2}$, work out the action of $R_x(\theta)$ then $R_y(\theta)$ on the vector $\mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Then work out the action of $R_y(\theta)$ then $R_x(\theta)$ on the same initial vector. In symbols, show that $R_yR_x\mathbf{r} \neq R_xR_y\mathbf{r}$. Make diagrams of these two different operations by drawing, on two (x,y,z) graphs, the initial vector, $\mathbf{r}$, the vector after applying the first rotation, and the final vector.

5. Rotations of spinors

Now that we're expert at rotating regular, good old vectors in 2D and 3D let's consider the more abstract problem of rotating an electron. An electron, as we've learned, is described in the quantum theory by a two-dimensional (not the same 2D plane where we rotated the good old vectors of the last section) sort of vector called a *spinor*. We need to figure out what mathematical **operator** we need to apply to the spinor to rotate it.

[Aside: This is an exact analog to what we did earlier in the chapter. There, the operator was the matrix $R(\boldsymbol{\theta})$ and we applied it to vectors, $\mathbf{r}$ in which 'inhabit' the 3D **vector space** which is often referred to as $\mathbb{R}^3$, the **Euclidean** space of three dimensions.]

We already know most of the mathematics necessary to rotate spinors – we learned it in sections 1–3. We'll walk through the process of rotating a spinor by doing some exercises.

When we measure the spin of an electron or a quark[4] – in fact, any spin–$\frac{1}{2}$ particle – we always find that it has only two values: either $+\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$. The constant,

$$\begin{aligned}\hbar = \frac{h}{2\pi} &= 1.054\,571\,68(18) \times 10^{-34}\ \text{J s} \\ &= 6.582\,119\,15(56) \times 10^{-22}\ \text{MeV s} \\ \hbar c &= 197.326\,968(17)\ \text{MeV fm},\end{aligned} \tag{1.85}$$

is called *Planck's constant*. It is a measure of the 'granularity' of the quantum world. As an example, if you had a photon, γ (a 'quantum' of the electromagnetic field) of frequency ν, then its energy is

$$E_\gamma = h\nu \tag{1.86}$$

and only that amount of energy (for this frequency) can be exchanged between the electromagnetic field and something else (like an electron, for example). These numbers for $\hbar$, no matter what conventional units that you use to describe them, are indecipherably small. That's why the quantum physics is usually[5] only important when we deal with submicroscopic objects like electrons and quarks.

Notice that $\hbar$ has the units of angular momentum:

$$\begin{aligned}[\hbar] &= \text{J s} = \text{kg m}^2\ \text{s}^{-1} \\ &= \text{m kg}\ \frac{\text{m}}{\text{s}} \\ &= [\mathbf{r}][\text{mass}][\mathbf{v}] \\ &= [\mathbf{r} \times \mathbf{p}]\end{aligned} \tag{1.87}$$

([ ] means "units of") and we recognize $\mathbf{r} \times \mathbf{p}$ as the orbital angular momentum $\mathbf{L}$. (From here on out we shall assume that we are using a set of units where $\hbar = c = 1$ and so neither $\hbar$ nor c will appear in our formulas.)

---

[4] One way to measure the intrinsic spin of a particle is to subject it to an external magnetic field and observe, for example, the Zeeman effect.

[5] There are some cases where macroscopic objects can exhibit quantum behavior. For example, a several ton aluminum cylinder, suspended in a vat of liquid helium, oscillates like a quantum mechanical object. This could be used as a gravitational wave detector.

(When we measure the spin of the electron, we're not measuring **L**. The orbital angular momentum is the angular momentum of an electron as it moves around some point in space. The intrinsic angular momentum – the "spin" – is the angular momentum that the electron has because it looks like it's spinning – but it's not *really*? Confusing, isn't it?)

Turning now to the spinor description of the electron, if we know that electrons are described mathematically by spinors and we know that measuring an electron's spin can result in only one of two possible values, $+\frac{1}{2}, -\frac{1}{2}$ then which spinors describes an electron with spin–'up' or 'down' along the $z-$axis? You might guess that they're given by:

$$|+\tfrac{1}{2}\rangle = |\uparrow\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \; |-\tfrac{1}{2}\rangle = |\downarrow\rangle = \begin{pmatrix}0\\1\end{pmatrix}. \tag{1.88}$$

These spinors are **eigenvectors** ('eigen' means 'characteristic' in German) of the Pauli matrix, σ_z which means that when you operate on the spinor with σ_z you get the same spinor back (the components are the same) times some number as follows:

$$\begin{aligned}\sigma_z|\uparrow\rangle &= (+1)|\uparrow\rangle \\ \sigma_z|\downarrow\rangle &= (-1)|\downarrow\rangle.\end{aligned} \tag{1.89}$$

Exercise 1.41. Verify these equations using the expression for the Pauli matrices, Eqs.(1.43).

Those numbers on the right-hand-side, ± 1, are the spins of the electron. "But," you think, "you said that the spin could have one of two values, $+\frac{1}{2}$ or $-\frac{1}{2}$. Not ± 1." That's because the spin is actually defined as:

$$\mathbf{S} = \frac{1}{2}\boldsymbol{\sigma} \tag{1.90}$$

$$\mathbf{S} = (S_x, S_y, S_z). \tag{1.91}$$

Incidentally, we've introduced a new notation here, the **ket**, $|\ \rangle$. In the quantum theory, systems are described by **quantum states** or just 'states'. The *ket* is a generic symbol for writing a 'state' of a system – you can put just about any symbol inside the ket, like $|\uparrow\rangle$ or $|\psi\rangle$, to denote different states. And a spinor describes the **spin state** of a spin-$\frac{1}{2}$ particle, like an electron. The ket has a natural 'partner' called a **bra**, written $\langle\ |$. For spinors, the *bra* is the Hermitian conjugate of the ket.

$$|\psi\rangle^\dagger = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}^\dagger = (\psi_1^*, \psi_2^*) = \langle\psi|. \tag{1.92}$$

Soon, we'll see where the weird names 'bra' and 'ket' come from.

Now, we're in a position to ask a more general question. Suppose a general spinor is given as

$$\begin{aligned}\psi &= \begin{pmatrix}\psi_\uparrow\\ \psi_\downarrow\end{pmatrix} \\ &= \psi_\uparrow \begin{pmatrix}1\\0\end{pmatrix} + \psi_\downarrow \begin{pmatrix}0\\1\end{pmatrix} \\ &= \psi_\uparrow|\uparrow\rangle + \psi_\downarrow|\downarrow\rangle\end{aligned} \tag{1.93}$$

Now, this is not an eigenstate of σ_z or S_z.

Exercise 1.42. Prove thate Eq.(1.93) is not an eigenstate of σ_z.

Therefore, the electron described by this spinor doesn't have a definite spin state. We'll see this a lot: if a particle's state is not described by an eigenvector of some operator (like S_z) then we're not guaranteed to get a particular value when we measure the quantity corresponding to that operator.

In fact, the situation that we have in the case of the spin of an electron described by the spinor of Eq.(1.93) is that when we measure it's spin along the $z-$direction we can get *either* $+\frac{1}{2}$ *or* $-\frac{1}{2}$!!! This is quantum weirdness at its best. The state of an electron is described by a mathematical object that gives you either possible value of the spin along the $z-$direction. If you know the story of Schrödinger's cat, you should be thinking of that beast's uncertain fate right now.

There is something we can determine though, from the spinor of Eq.(1.93). We can figure out the relative probability that a measurement gives you the values $+\frac{1}{2}$ or $-\frac{1}{2}$ from it. Here's how.

Consider a general spinor, ψ. It can be represented, as we've seen, by a complex column vector of which we can take the Hermitian conjugate (complex conjugate then transpose):

$$\psi = |\psi\rangle = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} \qquad \psi^\dagger = \langle\psi| = (\psi_1^*, \psi_2^*)\,. \tag{1.94}$$

If we have two general spinors χ and ψ then we define their **overlap** as:

$$\langle\chi|\psi\rangle = \chi^\dagger\psi \tag{1.95}$$

$$= (\chi_1^*, \chi_2^*)\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} \tag{1.96}$$

$$= \chi_1^*\psi_1 + \chi_2^*\psi_2. \tag{1.97}$$

This is just the generalization of the **dot** or **scalar** product for regular vectors, $\mathbf{r}_1 \cdot \mathbf{r}_2$ which has the geometric interpretation of being the projection of $\mathbf{r}_1$ along $\mathbf{r}_2$ (or vice versa). So the **overlap** is the "projection of χ onto ψ." Here, too, we see the origin of the names 'bra' and 'ket' – when they are sandwiched together like this $\langle\ \ |\ \ \rangle$ they look like a 'bracket'. This terminology is due to Paul Adrienne Maurice Dirac, one of the great physicists of the 20th century and a founder of the quantum theory.

Now, it's simple to see the answer to our question: What is the probability that a given measurement of the spin along the $z-$direction of an electron described by the spinor ψ gives $\pm\frac{1}{2}$? It must be related to the *projections*:

$$\langle\uparrow|\psi\rangle = \langle\uparrow|\big(\psi_\uparrow|\uparrow\rangle + \psi_\downarrow|\downarrow\rangle\big) = \psi_\uparrow \tag{1.98}$$

$$\langle\downarrow|\psi\rangle = \langle\downarrow|\big(\psi_\uparrow|\uparrow\rangle + \psi_\downarrow|\downarrow\rangle\big) = \psi_\downarrow, \tag{1.99}$$

where we've used the facts that

$$\langle\uparrow|\uparrow\rangle = \langle\downarrow|\downarrow\rangle = 1 \tag{1.100}$$

$$\langle\uparrow|\downarrow\rangle = \langle\uparrow|\downarrow\rangle = 0. \tag{1.101}$$

Exercise 1.43. Prove these equations from the definitions of $|\uparrow\rangle$ and $|\downarrow\rangle$ in Eq.(1.88)

But $\psi_\uparrow$ and $\psi_\downarrow$ can't directly be the probabilities. Why not? Think about it. A probability is the chance that something happens. "Four out of five dentists." This ratio, 0.8, is a real number. But what are $\psi_\uparrow$ and $\psi_\downarrow$? Are they real?

EXERCISE 1.44. What sort of numbers are $\psi_\uparrow$ and $\psi_\downarrow$?

In fact, it turns out, and it's not at all obvious that it turns out this way – you'll just have to take Max Born's word for it (Born was a 20th century German mathematician and physicist who worked on the foundations of the quantum theory for which he won the Nobel prize in physics in 1954), that the probabilities for getting spin–up or spin–down answers from ψ, $\mathcal{P}_\uparrow(\psi)$ or $\mathcal{P}_\downarrow(\psi)$ are

$$\mathcal{P}_\uparrow(\psi) = |\langle\uparrow|\psi\rangle|^2 = |\psi_\uparrow|^2 \tag{1.102}$$

$$\mathcal{P}_\downarrow(\psi) = |\langle\downarrow|\psi\rangle|^2 = |\psi_\downarrow|^2 . \tag{1.103}$$

EXERCISE 1.45. Suppose $\langle\psi|\psi\rangle = 1$. Calculate the sum $\mathcal{P}_\uparrow(\psi) + \mathcal{P}_\downarrow(\psi)$. {Hint: write out the $\langle\psi|\psi\rangle$ in terms of the components $\psi_\uparrow$ and $\psi_\downarrow$.}

This problem shows that, if we **normalize** the state $|\psi\rangle$ to unity (or one), which means $\langle\psi|\psi\rangle = 1$, then the probabilities of having a spin–up(down) electron, $\mathcal{P}_{\uparrow(\downarrow)}$ summed together exhaust the possible outcomes for measuring the spin. By adding up to one we say that, "The spin of the electron is either 'up' or 'down' 100% of the time." There are no other possibilities.

This constitutes the entire solution to the problem of how to describe the state of an electron with a general spinor with respect to a given **quantization axis**. Here we've used the $z-$axis as the quantization axis. That is why σ_z is a diagonal matrix. It corresponds to the spin along the quantization axis. Be careful though. When we want to ask question about the probability of finding the spin along a different axis, like say the $x-$axis, we have to take projections along this axis, not simply the eigenstates $|\uparrow\rangle$ and $|\downarrow\rangle$.

EXERCISE 1.46. Consider the state described by the ket $|\psi\rangle = \frac{1}{\sqrt{2}}|\uparrow\rangle + \frac{1}{\sqrt{2}}|\downarrow\rangle$. What is the probability that this state has spin-up and spin-down, $\mathcal{P}_\uparrow$ and $\mathcal{P}_\downarrow$?

For future problems we'll need to use the **raising operator**, σ_+ and **lowering operator**, σ_-. These are defined as

$$\sigma_+ = \frac{1}{2}(\sigma_x + i\sigma_y) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{1.104}$$

$$\sigma_- = \frac{1}{2}(\sigma_x - i\sigma_y) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} . \tag{1.105}$$

EXERCISE 1.47. Show that $\sigma_+|\uparrow\rangle = 0$, $\sigma_+|\downarrow\rangle = |\uparrow\rangle$, $\sigma_-|\uparrow\rangle = |\downarrow\rangle$, $\sigma_-|\downarrow\rangle = 0$

EXERCISE 1.48. Show that the state $|\psi\rangle$ in Prob. 13b is an eigenstate of σ_x [Eq.(1.43)]. Do this in two ways: i) Show that $|\psi\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ and use the matrix form for σ_x; ii) Solve for the σ_x in terms of $\sigma_\pm$ using Eq.(1.104) and Eq.(1.105) and then calculate $\sigma_x|\psi\rangle$ using the relations derived in Prob. 13c

Now we turn to the definition of a thing called a **matrix element**. If we have a general spinor $|\psi\rangle$ and some operator, A then when we act upon $|\psi\rangle$ with A from the left, $A|\psi\rangle$ we get some other state (in general), call it $|\chi\rangle = A|\psi\rangle$. Now, suppose there is a third, general state, $|\phi\rangle$. We now can take the overlap of $|\phi\rangle$ with $|\chi\rangle$. (Remember that we need to take the Hermitian conjugate of one of them to take the overlap.) We have

$$\mathcal{M} = \langle\phi|\chi\rangle = \langle\phi|A|\psi\rangle. \tag{1.106}$$

Since the matrix element $\mathcal{M}$ is the overlap of a bra with a ket it must be a complex number, $\mathcal{M} \in \mathbb{C}$. Since $\mathcal{M}$ is complex, we can take it's complex conjugate:

$$\mathcal{M}^* = \langle\phi|A|\psi\rangle^* = \langle\psi|A^\dagger|\phi\rangle \tag{1.107}$$

where now, we have the order changed – ϕ, which was in the bra is now in the ket, and *vice versa*. Also, notice that we have $A^\dagger$ instead of A – we have taken the Hermitian conjugate of the operator A.

We also need to define the concept of the **expectation value** of an operator. This is the value that we expect for a given physical quantity if we measure it in a state described by ψ. For example, for the spin along the $z-$direction:

$$\langle\hat{\sigma_z}\rangle = \langle\psi|\sigma_z|\psi\rangle. \tag{1.108}$$

This is just a special case of a matrix element where the state in the ket is the same as the state in the bra.

Exercise 1.49. Prove that the expectation value $\langle\hat{H}\rangle$ of a Hermitian operator, $H = H^\dagger$ is real. You can do this by showing that the complex conjugate of the expectation value is equal to the expectation value $\langle\hat{H}\rangle^* = \langle\hat{H}\rangle$. To do this, write out the components of $\psi = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}$ and then take the complex conjugate and transpose of the entire expression. (Transpose of a real number does nothing to it.) Remember that $(AB)^T = B^T A^T$ (prove this) for matrices A and B

Exercise 1.50. Prove that:

$$\langle\hat{\sigma_z}\rangle = \langle\psi|\sigma_z|\psi\rangle = \psi_1^*\psi_1 - \psi_2^*\psi_2, \tag{1.109}$$

for a general state ψ.

Now, we're in a position to rotate the spinors. The elements of $SU(2)$ are the rotation operators for spinors. And when we apply an element of $SU(2)$ to a general spinor, ψ we get another spinor, ψ' which is rotated with respect to the original ψ. (Compare this to the 2D and 3D cases. They're exactly analogous.) We have

$$\begin{aligned}\psi' &= U(\boldsymbol{\theta})\psi \\ \psi = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} &\qquad \psi' = \begin{pmatrix}\psi_1'\\ \psi_2'\end{pmatrix} \\ U(\boldsymbol{\theta}) = e^{-i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}/2} &= e^{-i\mathbf{S}\cdot\boldsymbol{\theta}} = \cos\theta/2 - i\boldsymbol{\sigma}\cdot\hat{\boldsymbol{\theta}}\sin\theta/2.\end{aligned} \tag{1.110}$$

Compare the expression for $U(\boldsymbol{\theta})$ above with Eq.(1.56). There is a difference of a factor of $-\frac{1}{2}$ in the angles between them. We need to include this factor in order to get the rotation properties of the spinors correct.

Exercise 1.51. Show that $U(\boldsymbol{\theta})$ can be written as:

$$U(\boldsymbol{\theta}) = \begin{pmatrix}\cos\theta/2 - i\hat{\theta}_z\sin\theta/2 & (-i\hat{\theta}_x - \hat{\theta}_y)\sin\theta/2\\ (-i\hat{\theta}_x + \hat{\theta}_y)\sin\theta/2 & \cos\theta/2 + i\hat{\theta}_z\sin\theta/2\end{pmatrix}, \tag{1.111}$$

where $\boldsymbol{\theta} = \theta_x\hat{\mathbf{x}} + \theta_y\hat{\mathbf{y}} + \theta_z\hat{\mathbf{z}}$ and $\hat{\theta}_{x,y,z} = \theta_{x,y,z}/\theta$ where $\theta = \sqrt{\theta_x^2 + \theta_y^2 + \theta_z^2}$.

Exercise 1.52. Let $\hat{\theta}_x = \hat{\theta}_z = 0$. Then $\boldsymbol{\theta} = \theta\hat{\mathbf{y}}$ and $\hat{\theta}_y = 1$. Calculate the matrix $U(\boldsymbol{\theta} = \theta\hat{\mathbf{y}})$

EXERCISE 1.53. Calculate the spinor ψ' which results by applying this $U(\theta\hat{\mathbf{y}})$ to $\psi = |\uparrow\rangle$.

EXERCISE 1.54. Calculate the probabilities $\mathcal{P}_\uparrow(\psi')$ and $\mathcal{P}_\downarrow(\psi')$ for the rotated spinor.

EXERCISE 1.55. Calculate the expectation values $\langle\hat{\sigma_z}\rangle = \langle\psi'|\sigma_z|\psi'\rangle$, $\langle\hat{\sigma_x}\rangle = \langle\psi'|\sigma_x|\psi'\rangle$. and $\langle\hat{\sigma_y}\rangle = \langle\psi'|\sigma_y|\psi'\rangle$. Show that they are equal to $\langle\hat{\sigma_z}\rangle = \cos\theta$, $\langle\hat{\sigma_x}\rangle = \sin\theta$ and $\langle\hat{\sigma_y}\rangle = 0$.

EXERCISE 1.56. What angle θ do we need to rotate the spinor $\psi = |\uparrow\rangle$ about the y–axis (as we have just done in Prob. 16b) so that the new state, $|\psi'\rangle$ is the same state in Prob. 13b?

EXERCISE 1.57. Explain why this angle θ makes sense physically.

EXERCISE 1.58. Draw a simple diagram showing the initial spin along the $z-$axis, the final spin along its axis, and the rotation axis and angle.

EXERCISE 1.59. Use the results of Prob. 18 to determine the values of $\langle\sigma_{\hat{x,y,z}}\rangle$ for this value of θ? Explain why these values make sense by looking at the final spin of the electron in your simple diagram.

CHAPTER 2

Matrix representation

Before we go on to the problem of an electron in a magnetic field we will consider the problem of the diagonalization of a matrix for which we often have the occasion in the course of doing quantum mechanics. Diagonalization yields the eigenvectors and eigenvalues of Hermitian operators. Although we have already used some matrix representation of operators (like the Pauli $\boldsymbol{\sigma}$ matrices) and states (like the spinor $\begin{pmatrix}1\\0\end{pmatrix}$ for the $|\uparrow\rangle$ state), this chapter seeks to systematize and generalize their use. The key notion in this chapter is the expression of a general Hermitian operator in its *matrix representation*.

1. Matrix representation of operators

Consider a Hermitian operator $\hat{H}$, $\hat{H}^\dagger = \hat{H}$. Let's calculate its matrix elements with respect to a set of **orthonormal** states $\{|a\rangle\}_{a=1}^n$ or **basis**. (The notation $\{|a\rangle\}_{a=1}^n$ where the label a runs over the integers $1,\ldots,n$ means 'the set of n kets.') The matrix elements are

$$H_{ab} = \langle a|\hat{H}|b\rangle, \tag{2.1}$$

where $a = 1,\ldots,n$ and $b = 1,\ldots,n$ and "orthonormal" means the states satisfy

$$\langle a|b\rangle = \delta_{ab}, \tag{2.2}$$

where δ_{ab} is the familiar Kronecker δ function, $\delta_{ab} = 0, 1$ for $a = b, a \neq b$, respectively. The operator may then be represented by the following matrix

$$\hat{H} = \begin{pmatrix} H_{11} & H_{12} & \ldots & H_{1n} \\ H_{21} & H_{22} & \ldots & H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{n1} & H_{n2} & \ldots & H_{nn} \end{pmatrix}. \tag{2.3}$$

There are n^2 complex numbers, H_{ab} giving the above form for the rank–n (square) matrix.

The Hermiticity $\hat{H}^\dagger = \hat{H}$ condition implies

$$\begin{aligned} H_{ba} &= \langle b|\hat{H}|a\rangle = \langle a|\hat{H}^\dagger|b\rangle^* \\ &= \langle a|\hat{H}|b\rangle^* = H_{ab}^*. \end{aligned} \tag{2.4}$$

Now, our objective is to diagonalize this matrix. When we do, we'll end up with one that looks like

$$\hat{h} = \begin{pmatrix} h_1 & 0 & \dots & 0 \\ 0 & h_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h_n \end{pmatrix}, \tag{2.5}$$

with zeros everywhere but along the diagonal. In order to do this we need to find the eigenvectors and eigenvalues of this matrix. There are n each of them and they satisfy

$$\hat{H}|\alpha\rangle = h_\alpha|\alpha\rangle. \tag{2.6}$$

The n eigenvectors of $\hat{H}$ are $\{|\alpha\rangle\}_{\alpha=1}^n$ and we denote them as $|\alpha\rangle$ in contrast with $|a\rangle$. They too are orthonormal. The n real numbers (real since $\hat{H}$ is Hermitian) h_α are the eigenvalues of the operator $\hat{H}$, each one corresponding to the particular $|\alpha\rangle$ with the same α.

Now we can see what it means to diagonalize a matrix. We are *changing the basis* to which we refer when we calculate the matrix elements of the Hermitian operator. To see this, note that in the "a-basis" we calculate the general matrix element as in Eq.(2.1). We contrast the evaluation in the a-basis with that in the α-basis:

$$H_{ab} = \langle a|\hat{H}|b\rangle \tag{2.7}$$

$$\begin{aligned} H_{\alpha\beta} &= \langle\alpha|\hat{H}|\beta\rangle \\ &= \langle\alpha|h_\beta|\beta\rangle \\ &= h_\beta\langle\alpha|\beta\rangle \\ &= h_\beta\delta_{\alpha\beta} \\ &= h_\alpha\delta_{\alpha\beta}, \end{aligned} \tag{2.8}$$

where we used Eq.(2.6) to get from the first line to the second line. And we used the fact that the $|\alpha\rangle$ are orthonormal to get from the third to fourth lines.

So you see that if the off-diagonal components of $\hat{H}$ are non-zero, we're in the a-basis. If the off-diagonal matrix elements are *all* zero, then we're in the α-basis.

The situation that we usually face is: we know the matrix elements in the a-basis and we want to find the eigenvectors (the $|\alpha\rangle$'s) and eigenvalues (the h_α's).

2. Secular equation

Now, you might remember that to diagonalize a matrix you solve the equation

$$\left|\hat{H} - hI\right| = 0 \tag{2.9}$$

where the vertical bars mean "take the determinant" and I is the identity matrix. This is called the **secular equation**. The h here is just a real variable at this stage, and shouldn't be confused as the eigenvalues.

This secular equation results from the following considerations. Suppose we rewrite Eq.(2.6) as

$$[\hat{H} - h_\alpha]|\alpha\rangle = 0. \tag{2.10}$$

We want to convert this *operator* equation into a *matrix* equation. That is, we want to express it in a **matrix representation**. In order to do this we need to express everything – $\hat{H}$ and $|\alpha\rangle$ – as matrices. The basis that we're going to work in is the a-basis, of course.

(Actually there are an infinite number of bases that one can work with. But there's only one[1] basis which diagonalizes $\hat{H}$.)

In order to do this we might think as follows: the ket

$$|\xi\rangle = \hat{H}|\alpha\rangle \tag{2.11}$$

which appears on the left-hand-side of Eqs.(2.6),(2.10) and can be expressed in the a-basis simply by taking the overlap

$$\langle a|\xi\rangle = \langle a|\hat{H}|\alpha\rangle \equiv H_{a\alpha}. \tag{2.12}$$

But we've only done '*half*' the job. We need both subscripts on $\hat{H}$ to be in the a-basis. If we could manage to squeeze something like $|a\rangle\langle a|$ in between the $\hat{H}$ and the $|\alpha\rangle$ we'll be in business.

Exercise 2.1. Assume that we can write

$$\sum_{a=1}^{n} |a\rangle c_a = 1, \tag{2.13}$$

here we also assume that $|a\rangle$ and c_a don't commute. Using the fact that the a-basis is orthonormal $\langle a|b\rangle = \delta_{ab}$ show that

$$c_a = \langle a| \tag{2.14}$$

so that

$$\sum_{a=1}^{n} |a\rangle\langle a| = 1. \tag{2.15}$$

Eq.(2.15) is called **completeness** and it's of central importance in the quantum theory.

We can proceed with with our mission – to express $\hat{H}$ and $|\alpha\rangle$ as matrices – using the completeness relation. We write

$$\begin{aligned}
\langle a|\xi\rangle &= \langle a|\hat{H}|\alpha\rangle \\
&= \langle a|\hat{H}\cdot 1\cdot|\alpha\rangle \\
&= \langle a|\hat{H}\cdot\sum_{b=1}^{n}|b\rangle\langle b|\cdot|\alpha\rangle \\
&= \sum_{b=1}^{n}\langle a|\hat{H}|b\rangle\langle b|\alpha\rangle \\
\xi_a &= \sum_{b=1}^{n} H_{ab}\psi_b,
\end{aligned} \tag{2.16}$$

[1]We're assuming that no two of the eigenvalues h_α are the same – *ie.* that the **spectrum** of $\hat{H}$ in **non-degenerate**.

where we have written

$$\xi_a = \langle a|\xi\rangle \tag{2.17}$$

$$\psi_b = \langle b|\alpha\rangle \tag{2.18}$$

as the *components* of $|\xi\rangle, |\alpha\rangle$ in the a-basis. You see that by "inserting a complete set of states"[2] at the third line above we have converted $\hat{H}$ in the "mixed representation," $H_{a\alpha}$ to the matrix representation in the a-basis. And we have gotten, as a bonus, the matrix representation of $|\alpha\rangle$ (it's an n-by-1 matrix, but still a matrix).

EXERCISE 2.2. Use completeness of the a-basis to write Eq.(2.10) in its matrix representation.

Now, using the results of Ex.(2.2) we can obtain the secular equation as follows. From the result of Ex.(2.2) we have

$$\sum_b (H_{ab} - h_\alpha \delta_{ab})\psi_b = 0. \tag{2.19}$$

Now, suppose we could invert the matrix $M = \hat{H} - hI$ on the left-hand-side of this equation. Then we could write

$$\psi_a = \sum_b M^{-1}_{ab} \cdot 0 \tag{2.20}$$

which would mean that all the ψ_a are zero – not a very interesting solution to the above equations. Therefore, in order to have non-trivial solutions, M must be **singular**, *ie.* not invertible. This means it's determinant is zero.

$$\det M = \left|\hat{H} - hI\right| = 0, \tag{2.21}$$

which is the secular equation.

3. Completeness

Completeness –

$$\sum_{a=1}^{n} |a\rangle\langle a| = 1 \tag{2.22}$$

– is such an important property of quantum systems that we should take a few moments to understand its content.

Any state, for example $|\zeta\rangle$, can be expressed in the matrix representation by virtue of the completeness relation

$$\begin{aligned} |\zeta\rangle &= 1 \cdot |\zeta\rangle \\ &= \sum_a |a\rangle\langle a| \cdot |\zeta\rangle \\ &= \sum_a |a\rangle\langle a|\zeta\rangle \\ &= \sum_a |a\rangle\zeta_a. \end{aligned} \tag{2.23}$$

(We won't always write the range over which a runs in the summation symbol $\sum$, but we assume it's from 1 to n.)

[2]This has been referred to by, at least, R.V. Reid of UC-Davis as "The old German trick."

Suppose I define an *operator*

$$P_a = |a\rangle\langle a|. \tag{2.24}$$

What is it?

EXERCISE 2.3. Show that the name for P_a of **projection operator** makes sense by acting on $|\zeta\rangle$ with it.

So we see that the projection operator isolates or projects-out the component of a general state, like ζ, "along" $|a\rangle$. And we also see that if we sum the projection operators along each state in the a-basis

$$\sum_a P_a|\zeta\rangle = \sum_a |a\rangle\zeta_a = |\zeta\rangle, \tag{2.25}$$

so that

$$\sum_a P_a = 1. \tag{2.26}$$

EXERCISE 2.4. Show that the projection operators obey the following relations:

i) $P_a P_b = 0$, $(a \neq b)$;
ii) $P_a P_a = P_a$.

(Or, compactly, $P_a P_b = \delta_{ab} P_a$.)

4. Rank–2 example – Two-level system

Let's see how this all works for a rank–2 matrix. Using the definition Eq.(2.3) with $n = 2$ we have

$$\hat{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \qquad\qquad hI = \begin{pmatrix} h & 0 \\ 0 & h \end{pmatrix}. \tag{2.27}$$

The secular equation, Eq.(2.9) gives:

$$\left|\hat{H} - hI\right| = \begin{vmatrix} H_{11} - h & H_{12} \\ H_{21} & H_{22} - h \end{vmatrix} = 0,$$

$$h^2 - h(H_{11} + H_{22}) + H_{11}H_{22} - |H_{12}|^2 = 0. \tag{2.28}$$

So we see that the secular equation is just a second order polynomial in h. It therefore has 2 real roots. And these roots are the eigenvalues.

EXERCISE 2.5. Derive Eq.(2.28) and complete the square to solve the quadratic equation in h for the two roots, h_1 and h_2.

EXERCISE 2.6. Write the result of Ex.(2.5) in terms of the trace and determinant of the matrix $\hat{H}$.

Now that we know the eigenvalues, h_1, h_2 we can find the corresponding eigenvectors. The eigenvector equations, Eq.(2.6) are

$$\hat{H}|\alpha_1\rangle = h_1|\alpha_1\rangle \tag{2.29}$$

$$\hat{H}|\alpha_2\rangle = h_2|\alpha_2\rangle, \tag{2.30}$$

where we have written the states $|\alpha\rangle$ with $\alpha = 1$ and $\alpha = 2$ in Eq.(2.6) as α_1 and α_2 to make clear that we're referring to the eigenvectors. Either of these equations can be written in terms of the the basis of $\{a\}_{a=1}^{2} \equiv \{a_1, a_2\}$ as

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = h \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \tag{2.31}$$

which can be written:

$$H_{11}\psi_1 + H_{12}\psi_2 = h\psi_1 \tag{2.32}$$

$$H_{21}\psi_1 + H_{22}\psi_2 = h\psi_2. \tag{2.33}$$

Here $\psi_1 = \langle a_1|\alpha\rangle$ and $\psi_2 = \langle a_2|\alpha\rangle$ for $\alpha = 1$ or $\alpha = 2$ depending on which eigenstate we're talking about.

EXERCISE 2.7. Show that if we write the solution for either state $|\alpha_1\rangle$ or $|\alpha_2\rangle$ as

$$|\alpha\rangle = \frac{1}{N} \begin{pmatrix} \sqrt{\frac{H_{12}}{h - H_{11}}} \\ \sqrt{\frac{H_{21}}{h - H_{22}}} \end{pmatrix}, \tag{2.34}$$

where N is given by

$$N = \sqrt{\frac{2h - (H_{11} + H_{22})}{|H_{12}|}}, \tag{2.35}$$

then we obtain the secular equation when we substitute this solution for $|\psi\rangle$ into either of the equations (2.32) or (2.33). (Note that the constant, $\frac{1}{N}$ drops out when $|\alpha\rangle$ is substituted into Eqs.(2.32),(2.33) so you don't need to write it out.)

EXERCISE 2.8. Write down the explicit forms for $|\alpha_1\rangle$ and $|\alpha_2\rangle$. Assume that $h_1 < h_2$. (You need to think about what quantity you should replace h with in Eqs.(2.34) and (2.35)).

EXERCISE 2.9. Show that the normalization of the above states $|\psi\rangle$ is one:

$$\langle\psi|\psi\rangle = 1, \tag{2.36}$$

by using the secular equation which can be written $(h - H_{11})(h - H_{22}) = |H_{12}|^2$.

EXERCISE 2.10. **Extra Credit**: Show that the eigenvectors are orthogonal $\langle 1|2\rangle = 0$.

5. Diagonalized spin matrix

Let's use this machinery to calculate the direction $\hat{\mathbf{n}}$ an electron's spin points in if it has the wave function

$$|\psi\rangle = \psi_\uparrow|\uparrow\rangle + \psi_\downarrow|\downarrow\rangle. \tag{2.37}$$

In the case that $\hat{\mathbf{n}} = \hat{\mathbf{z}}$ we have $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}} = \sigma_z$ and we know that the eigenstates $|\uparrow\rangle$ and $|\downarrow\rangle$ satisfy the eigenvalue equations

$$\begin{aligned} \sigma_z|\uparrow\rangle &= (+1)|\uparrow\rangle \\ \sigma_z|\downarrow\rangle &= (-1)|\downarrow\rangle \end{aligned} \tag{2.38}$$

The generalization of this equation to an arbitrary direction $\hat{\mathbf{n}}$ is

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}|\psi\rangle = s|\psi\rangle \tag{2.39}$$

where $s = \pm 1$ is the eigenvalue of spin operator $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$ along a direction $\hat{\mathbf{n}}$.

We can write the direction unit vector $\hat{\mathbf{n}}$ in terms of its components

$$\hat{\mathbf{n}} = (n_x, n_y, n_z) \tag{2.40}$$

$$\hat{\mathbf{n}} \cdot \hat{\mathbf{n}} = n_x^2 + n_y^2 + n_z^2 = 1. \tag{2.41}$$

EXERCISE 2.11. Write the operator $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$ as a rank–2 matrix in terms of n_x, n_y, n_z.

EXERCISE 2.12. Write Eq.(2.39) in terms of components, multiplying the matrix $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$ into the spinor $|\psi\rangle$.

EXERCISE 2.13. Find the roots of the secular equation

$$|\boldsymbol{\sigma} \cdot \hat{\mathbf{n}} - sI| = 0. \tag{2.42}$$

EXERCISE 2.14. Write the components of the vector $\hat{\mathbf{n}}$ in terms of the polar angle θ and the azimuthal angle ϕ (measured from the x–axis) as in Fig.(2).

EXERCISE 2.15. Show that the eigenvectors

$$|\alpha_1\rangle = \begin{pmatrix} \cos\frac{\theta}{2} e^{-i\phi/2} \\ \sin\frac{\theta}{2} e^{+i\phi/2} \end{pmatrix} \tag{2.43}$$

$$|\alpha_2\rangle = \begin{pmatrix} -\sin\frac{\theta}{2} e^{-i\phi/2} \\ \cos\frac{\theta}{2} e^{+i\phi/2} \end{pmatrix} \tag{2.44}$$

satisfy Eq.(2.39). (You will find useful the half-angle formulas $\sin\theta = 2\sin\theta/2\cos\theta/2$, $2\cos^2\theta/2 = 1 + \cos\theta$, and $2\sin^2\theta/2 = 1 - \cos\theta$.)

EXERCISE 2.16. Using Eqs.(2.43),(2.44) show that the eigenvectors $|\alpha_1\rangle$ and $|\alpha_2\rangle$ are orthogonal, *ie.* $\langle\alpha_1|\alpha_2\rangle = 0$ and that they are normalized to one.

EXERCISE 2.17. Show that the state you obtained by rotating $|\uparrow\rangle$ in Ex.(1.53) is the same state that you get when you set $\phi = 0$ in the result of Ex.(2.14).

CHAPTER 3

Spin–$\frac{1}{2}$ particle in a magnetic field

1. Introduction

Now we turn to the problem of finding the 'motion' of a single electron (by 'electron' we still mean any spin–$\frac{1}{2}$ particle) in a constant magnetic field.

In general, the magnetic field is described by a vector in ordinary three dimensional Euclidean space, $\mathbb{R}^3$. It's written as $\mathbf{B}$ and it can, in general depend on the position, $\mathbf{r}$:

$$\mathbf{B}(\mathbf{r}) = (B_x(\mathbf{r}), B_y(\mathbf{r}), B_z(\mathbf{r})) \tag{3.1}$$

$$= B_x(\mathbf{r})\hat{\mathbf{x}} + B_y(\mathbf{r})\hat{\mathbf{y}} + B_z(\mathbf{r})\hat{\mathbf{z}}. \tag{3.2}$$

Note that, in general, each component depends on the position. For this section we'll take the magnetic field to be

$$\mathbf{B}(\mathbf{r}) = (0, 0, B), \tag{3.3}$$

with $B > 0$ and independent of position.

Aside: If a charged particle (charge Q) *without spin* enters a region of space with a non-zero magnetic field, it experiences a force given by

$$\mathbf{F} = Q\mathbf{v} \times \mathbf{B}, \tag{3.4}$$

which is perpendicular to both its velocity $\mathbf{v}$ and $\mathbf{B}$. We aren't considering this kind of force in what follows. The force we are talking about is independent of the velocity (it is experienced even by particles at rest for which $\mathbf{F} = 0$) and only affects particles with intrinsic angular momentum, *ie.* spin.

By virtue of the fact that the electron has spin, it has what's called a *magnetic moment* which means that it acts like a tiny 'bar magnet.' Its energy therefore depends on the orientation of the spin with respect to the magnetic field. This energy is given by

$$\hat{H}_{spin} = -\boldsymbol{\mu} \cdot \mathbf{B} \tag{3.5}$$

where $\boldsymbol{\mu}$ is the magnetic moment. The magnetic moment is given in terms of the Pauli spin matrices

$$\boldsymbol{\mu} = \frac{e}{2m}\mathbf{S} = \frac{e}{2m}\frac{\boldsymbol{\sigma}}{2}, \tag{3.6}$$

$$\mathbf{S} = \frac{\boldsymbol{\sigma}}{2}, \tag{3.7}$$

where e is the magnitude of the electron charge, $e = 1.6 \times 10^{-19}$ C, where C means 'Coulombs' and m is the electron's mass which we usually write as an energy $m = 0.511$ MeV. (Special relativity shows that mass and energy are equivalent.) Here we see that the magnetic moment $\boldsymbol{\mu}$ is proportional to the *electron spin* $\mathbf{S}$ which is *half* of the Pauli spin matrices.

Now, forget for a moment that the electron is a quantum object whose spin is described by the matrices $\boldsymbol{\sigma}$ and just pretend that it's an ordinary vector

$$\boldsymbol{\mu} = (\mu_x, \mu_y, \mu_z), \tag{3.8}$$

and assume that the magnetic field is given by Eq.(3.3). Then the energy is

$$H_{spin} = -\mu B \cos\theta \tag{3.9}$$

where $\mu = |\boldsymbol{\mu}|$ and θ is the angle between the $\mathbf{B}$ field and $\boldsymbol{\mu}$ (θ is the *polar angle*).

When the spin is aligned (anti-aligned) with the magnetic field $\theta = 0(\pi/2)$ and the energy is

$$H_{spin} = \begin{cases} -\mu B & \theta = 0 \\ +\mu B & \theta = \pi/2 \end{cases} \tag{3.10}$$

so, clearly, the configuration where the spin is aligned is favored since it has lower energy.

EXERCISE 3.1. Plot H_{spin} as a function of θ.

EXERCISE 3.2. Differentiate H_{spin} with respect to θ.

EXERCISE 3.3. Calculate the **torque**, $\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}$ ($\times$ means the vector or cross product) using the $\mathbf{B}$ field given in Eq.(3.3). How is related to the result of Ex.(3.2)? This is a general result: The derivative of the energy with respect to a coordinate gives a generalized force (like a torque).

Let's return to the quantum theory now. The magnetic moment $\boldsymbol{\mu}$ is not described by a classical three vector – it's described by the Pauli spin matrices. So, for the $\mathbf{B}$ field of Eq.(3.3) we get:

$$\hat{H}_{spin} = -\frac{eB}{2m}\boldsymbol{\sigma}_z = -\frac{eB}{2m}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.11}$$

The 'motions' of a quantum system are described by the **time dependent Schrödinger equation** (TDSE). It is a differential equation that relates the **Hamiltonian** (or energy), Eq.(3.5) to a time derivative of the wave function $|\psi\rangle$:

$$i\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}_{spin}|\psi(t)\rangle. \tag{3.12}$$

In the present case of the electron in the magnetic field, the wave function is a constant (in ordinary three dimensional $\mathbb{R}^3$ space) spinor which depends on time:

$$|\psi(t)\rangle = \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix}. \tag{3.13}$$

The TDSE [Eq.(3.12)] in words says that if we act with the Hamiltonian, here $\hat{H}_{spin}$ then we get the (partial) time derivative of the wave function (up to the constant factor i). This is a general feature of equations-of-motion (EOM). They relate time derivatives of dynamical variables (here, $|\psi(t)\rangle$ is the dynamical variable) to functions of the dynamical variables. Think of Newton's second law:

$$\mathbf{F} = m\mathbf{a} \tag{3.14}$$

$$\mathbf{a} = \frac{d^2\mathbf{r}(t)}{dt^2} \tag{3.15}$$

$$\frac{d^2\mathbf{r}(t)}{dt^2} = \frac{1}{m}\mathbf{F}, \tag{3.16}$$

where m is the mass of the particle and $\mathbf{F}$ is the force acting on it. The last line shows that Newton's second law is an EOM – it relates, in this case, the *second* time derivative of the position of a particle – which is now the dynamical variable – to a function, which is a vector, $\mathbf{F}$.

2. Separation of variables

Returning to the TDSE – our first step is to solve for the time dependence of the wave function by the general method of solution of partial differential called **separation of variables**. We write the wave function as a product

$$|\psi(t)\rangle = T(t)|\psi(0)\rangle \tag{3.17}$$

$$|\psi\rangle \equiv |\psi(0)\rangle, \tag{3.18}$$

where $T(t)$ is a function of the time only. If we substitute this form into the SE we can write

$$\hat{H}_{spin}|\psi\rangle = \left[i\frac{1}{T(t)}T'(t)\right]|\psi\rangle, \tag{3.19}$$

where $T'(t) = \frac{dT(t)}{dt}$. Now, for a time-independent magnetic field $\mathbf{B}$ the $\hat{H}_{spin}$ is independent of the time. That means that only the factor in the square brackets $[\cdots]$ can depend on time. But since the left-hand-side (lhs) is independent of time in Eq.(3.19) then so must be the right-hand-side (rhs). So the factor in brackets is *a constant*:

$$\left[i\frac{1}{T(t)}T'(t)\right] = E \tag{3.20}$$

which we have arbitrarily (wink, wink) chosen to call E.

The solution to the equation

$$T'(t) = -iET(t) \tag{3.21}$$

is obvious if you know a little calculus. But it's also possible to explicitly derive the solution and the method by which we do this is frequently encountered in quantum mechanics and its generalization, quantum field theory, so I'll give an explanation of this *iterative solution* here.

3. Iterative solution

Integrate Eq.(3.21) with respect to t' from some initial time t_i to some final time t

$$T(t) = T(t_i) - iE\int_{t_i}^{t} dt' T(t'). \tag{3.22}$$

We've converted the differential equation to an integral equation. This is generally not thought of as a simplification because integral equations are more difficult to solve than differential ones. The rhs depends on the the function $T(t)$ that we're trying to find on the lhs. It looks like this is a nightmare.

Let's try something even more ridiculous looking – let's substitute the expression for $T(t')$ on the lhs into the rhs:

$$T(t) = T(t_i) - iE\int_{t_i}^{t} dt' \left[T(t_i) - iE\int_{t_i}^{t'} dt'' T(t'')\right]. \tag{3.23}$$

Note that in the factor in $[\cdots]$, the upper limit of integration is now t', *not* t, and that we have changed the integration variable (a 'dummy' variable) to t'' for this factor in order to consistently keep track of the *two separate* integrations over dt' and dt''. Moving things around, we can rewrite this as:

$$T(t) = T(t_i) - iE(t-t_i)T(t_i) + (-iE)^2 \int_{t_i}^{t} dt' \int_{t_i}^{t'} dt'' T(t''). \tag{3.24}$$

Now we see the utility in this iterative solution – we have obtained the second term $-iE(t-t_i)T(t_i)$ which is 'of order $\mathcal{O}(t-t_i)$', *ie.* it depends linearly on time. The next term we will get by substituting Eq.(3.22) into Eq.(3.24) will be second order, $\mathcal{O}(t-t_i)^2$, *ie.* quadratic in time.

EXERCISE 3.4. Prove Eq.(3.24).

EXERCISE 3.5. Show that the second order term is:

$$T(t_i)\frac{(-iE)^2}{2}(t-t_i)^2. \tag{3.25}$$

Carrying out the iterative procedure to all orders in $(t-t_i)$ we obtain

$$T(t) = T(t_i)\sum_{n=0}^{\infty}\frac{[-iE(t-t_i)]^n}{n!} \tag{3.26}$$

$$= T(t_i)e^{-iE(t-t_i)}. \tag{3.27}$$

EXERCISE 3.6. Show that $T(t_i = 0) = 1$.

Finally, we have arrived at the solution to the TDSE (setting $t_i = 0$)

$$\hat{H}_{spin}|\psi(t)\rangle = i\frac{\partial}{\partial t}|\psi(t)\rangle \tag{3.28}$$

$$|\psi(t)\rangle = e^{-iEt}|\psi\rangle \tag{3.29}$$

$$\hat{H}_{spin}|\psi\rangle = E|\psi\rangle \tag{3.30}$$

where the last equation is the time independent Scrhödinger equation (TISE), which we will now solve. (Actually, we've already solved it – the solutions are given by the kets $|1\rangle$ and $|2\rangle$ for a $\mathbf{B}$ field in a general direction, $\hat{\mathbf{n}}$.)

4. Time independent Schrödinger equation

Let's begin by assuming that the $\mathbf{B}$ field is given by Eq.(3.3). Then the Hamiltonian is

$$\hat{H}_{spin} = -\boldsymbol{\mu}\cdot\mathbf{B}, \tag{3.31}$$

$$= -\frac{eB}{2m}S_z, \tag{3.32}$$

$$S_z = \frac{\sigma_z}{2}, \tag{3.33}$$

and the TISE is

$$-\frac{eB}{2m}\frac{\sigma_z}{2}|\psi\rangle = E|\psi\rangle \tag{3.34}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix}\psi_1 \\ \psi_2\end{pmatrix} = -\frac{E}{\omega_L/2}\begin{pmatrix}\psi_1 \\ \psi_2\end{pmatrix} \tag{3.35}$$

$$\omega_L = \frac{eB}{2m} \tag{3.36}$$

where ω_L is the *Larmour frequency*.

EXERCISE 3.7. Prove that the units of ω_L are s^{-1}, a frequency. Use the units $[m]$ =MeV, $[eB]$ =MeV^2. "MeV" means "mega electron-volts" and it's the standard unit of mass or energy in nuclear physics. You'll have to put in a factor of $\hbar$ (which has units $[\hbar]$ =MeV s) which we have set to one.

We want to solve Eq.(3.35). Look at it – it looks just like Eq.(2.39) with $\hat{\mathbf{n}} = (0, 0, 1)$. Then we've already solved it! The eigenvalues are ± 1 and the eigenvectors are given in Eqs.(2.38).

EXERCISE 3.8. What are the eigenvalues of Eq.(3.35) as a function of ω_L?

EXERCISE 3.9. What are the eigenvectors?

If we write the eigenvectors as $|\psi_1\rangle$ and $|\psi_2\rangle$ and their corresponding eigenvalues E_1 and E_2 with $E_1 < E_2$ and form the *linear superposition*

$$|\psi(t)\rangle = c_1 e^{-iE_1 t}|\psi_1\rangle + c_2 e^{-iE_2 t}|\psi_2\rangle, \tag{3.37}$$

then this $|\psi(t)\rangle$ is a solution of the TDSE, Eq.(3.28). Here, c_1 and c_2 are general complex numbers.

EXERCISE 3.10. Prove this.

We now have determined the form for the general solution of the quantum EOM for a single spin in a constant magnetic field. What does the motion of the spin look like? In order to answer this question – as with any question regarding an experimentally determinable observable in quantum mechanics – we need to calculate the expectation values of some operator in some state.

EXERCISE 3.11. Calculate the expectation values of all three Pauli matrices, $\langle\hat{\sigma}_i\rangle = \langle\psi(t)|\sigma_i|\psi(t)\rangle, i = x, y, z$. Express your answer so that the fact that expectation values are real is explicitly demonstrated. Recall that $\Re z = \frac{z+z^*}{2}$ and $\Im z = \frac{z-z^*}{2i}$.

Solution:

$$\langle\hat{\sigma}_x\rangle = 2\Re(c_1^* c_2)\cos\omega_L t + 2\Im(c_1^* c_2)\sin\omega_L t \tag{3.38}$$

$$\langle\hat{\sigma}_y\rangle = -2\Re(c_1^* c_2)\sin\omega_L t + 2\Im(c_1^* c_2)\cos\omega_L t \tag{3.39}$$

$$\langle\hat{\sigma}_z\rangle = |c_1|^2 - |c_2|^2 . \tag{3.40}$$

These results demonstrate that the electron's spin *precesses* about the z–axis with frequency ω_L.

This is the principle at work in *nuclear magnetic resonance* (NMR) imaging. A sample, like a person's head, is subject to a magnetic field which sets all of the

nuclei which have non-zero spins to precessing. Then the sample is shot with a radio frequency (tens of megahertz) electromagnetic pulse. If the frequency of the pulse is resonant with (close to) some of the nuclear spin precession frequencies (which depend on the charge and mass of the nucleus) then the electromagnetic pulse will cause transitions of spins which are aligned with the magnetic field and while spins which are anti-aligned decay due to quantum and statistical fluctuations. By measuring the energy imparted to the spins and the energy emitted it is possible to obtain full three-dimensional imaging.

CHAPTER 4

Many-body quantum mechanics

So far, we have been talking about the quantum description and dynamics of a single particle (actually, just about its spin – at some point in the future you'll learn about the motion of particles in three dimensional configuration space). Now, we turn to the quantum description of two or more particles – many-body quantum mechanics.

For the purposes of this project, when we write down the wave function for two or three particles, we will *ignore interactions between them* – we treat the particles as *free* or non-interacting. We will study the interactions between the particles at a later stage, but we will ignore the effect of these interactions on the wave functions themselves.[1]

1. Two spin-$\frac{1}{2}$ particles

Suppose we have an electron (or a quark), call it electron "1" in a state described by the state $|\psi_1\rangle$ and another electron, call it "2" described by the state $|\chi_2\rangle$.[2] The subscripts denote which particle and the $\psi = \uparrow, \downarrow$ and $\chi = \uparrow, \downarrow$ denote the state the particle is in. There are four possible states that the two particles can be in and they're described by $|s_1 s_2\rangle$:

$$|\downarrow_1\downarrow_2\rangle, |\downarrow_1\uparrow_2\rangle, |\uparrow_1\downarrow_2\rangle, |\uparrow_1\uparrow_2\rangle, \tag{4.1}$$

for the possible values of $s_1 = \downarrow_1, \uparrow_1$ and $s_2 = \downarrow_2, \uparrow_2$. We'll call this basis the s-basis.[3] For now, it doesn't matter which order we write the s_1 and s_2 in – that is, $|s_1 s_2\rangle = |s_2 s_1\rangle$.

Exercise 4.1. Prove that we need 2^N states to describe the spin state of N electrons.

Exercise 4.2. Write down the states of three particles in the s-basis. Write them down in the following order: the first state has all spins $\downarrow$; for the rest of the states the rightmost spin flips back and forth each time; the second rightmost state flips every other time, etc. Look at the order in which I wrote down the states in Eq.(4.1) for guidance.

[1] You might ask yourself how we can get away with such a gross simplification, especially since QCD is the theory of *strong* interactions, and still say anything useful about the observed hadron data. For now, suffice it to say that we can learn some basic features of QCD within this simple-minded picture, but that we shouldn't content ourselves beyond the realization that this approach is anything other than a first step into a complex and beautiful but vast and challenging terrain.

[2] These states, $|\psi_1\rangle$ and $|\chi_2\rangle$ are just spinors.

[3] There are other bases we can use, as we'll see below.

Suppose we subject these electrons to a constant magnetic field like we did in the last section. Since these electrons aren't interacting the total energy of the two of them together is just

$$E = E_1 + E_2, \tag{4.2}$$

which means that the Hamiltonian operator $\hat{H}_{spin}$ is just the sum of the one-particle (or *one-body*) Hamiltonians

$$\hat{H}_{spin} = \hat{H}_{spin,1} + \hat{H}_{spin,2} \tag{4.3}$$

$$= -\frac{e}{2m}(\boldsymbol{\sigma}_1 \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \boldsymbol{\sigma}_2) \cdot \mathbf{B}, \tag{4.4}$$

where $\boldsymbol{\sigma}_{1,2}$ are the Pauli spin matrices for particles 1 and 2, respectively, $\otimes$ means **direct product**, and $\mathbb{1}_{1,2}$ are the rank-2 unit matrices for particles 1 and 2, respectively.

What are those $\otimes$'s and $\mathbb{1}$'s doing here? Suppose we're talking about a system with N electrons in it and all N electrons are described by $|\Psi\rangle$. The state $|\Psi\rangle$ depends on the spins of all the electrons. Then anytime we act on this state with an operator, $\mathcal{O}$ – *any* operator – this operator must act on the spins (more generically, the **coordinates**) of all N electrons – in more normal words, $\mathcal{O}$ must have 'instructions' for what do to *each and every* particle. So the term in Eq.(4.3) that's proportional to $\boldsymbol{\sigma}_1 \otimes \mathbb{1}_2$ means: "evaluate the spin of particle 1 and leave the spin of particle 2 alone."

So all operators, $\mathcal{O}$ are N-body operators, but we generally refer to operators in many-body systems by the number of particles that are affected by the operator. Then $\boldsymbol{\sigma}_1 \otimes \mathbb{1}_2$ is a 'one-body' operator and something like $\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2$ is a 'two-body' operator because it acts on the spins (coordinates) of both particle 1 and 2.

The symbol $\otimes$ means that the operators on either side act on different particles.[4] We will not usually write down the unit operators (matrices) $\mathbf{1}$ but they are understood to be there. Then we have

$$\hat{H}_{spin} = -\frac{e}{2m}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_2) \cdot \mathbf{B}. \tag{4.5}$$

Notice something about this Hamiltonian – if we interchange the labels 1 and 2 on the $\boldsymbol{\sigma}$'s, we get back the same operator. This is an axiom of quantum mechanics: *all particles of a given species are identical*. Electrons are indistinguishable. Protons are indistinguishable. Etc.[5] This means that the Hamiltonian must be the same no matter which way we label the particles. Another way of saying this is that "the Hamiltonian is **completely symmetric** under permutations of the particle labels (or indices)." This is called the 'principle of indistinguishability of particles' (PIP).

2. Principle of indistinguishability of particles (PIP)

Suppose we wish to describe a state where one of the particles is $\uparrow$ and the other is $\downarrow$. Which state should we use to describe it? Should it be $|\uparrow_1\downarrow_2\rangle$ or $|\downarrow_1\uparrow_2\rangle$? Seems like we have some trouble here.

[4] These particles are in different Hilbert spaces, $\mathcal{H}_1$ and $\mathcal{H}_2$ and the full Hilbert space is the **direct product** space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$.

[5] This is not so classically – given a labeling of the N *classical electrons* we can follow the trajectory of each in time and keep track of them forever. In quantum mechanics, we can't do this – as we'll see below. Also, this *axiom* can't be proven – other than experimentally, and there are plenty of experiments which verify it.

EXERCISE 4.3. Use the PIP to resolve this apparent ambiguity.

Resolving this ambiguity requires that we use *both* states $|\uparrow_1\downarrow_2\rangle$ or $|\downarrow_1\uparrow_2\rangle$. But how? The PIP not only constrains the form of the Hamiltonian, it also constrains the form of wave functions used to describe the system. Consider an operator P_{12} which acts on states to change the labels of the particles. That is:

$$P_{12}|\downarrow_1\downarrow_2\rangle = |\downarrow_2\downarrow_1\rangle = |\downarrow_1\downarrow_2\rangle \tag{4.6}$$
$$P_{12}|\downarrow_1\uparrow_2\rangle = |\downarrow_2\uparrow_1\rangle = |\uparrow_1\downarrow_2\rangle \tag{4.7}$$
$$P_{12}|\uparrow_1\downarrow_2\rangle = |\uparrow_2\downarrow_1\rangle = |\downarrow_1\uparrow_2\rangle \tag{4.8}$$
$$P_{12}|\uparrow_1\uparrow_2\rangle = |\uparrow_2\uparrow_1\rangle = |\uparrow_1\uparrow_2\rangle. \tag{4.9}$$

Before getting to the constraint that PIP places on the wave functions, let's have a look at these relations. We notice that if we write the spin of particle 1 first and then the spin of particle 2 second, we can drop the subscripts 1 and 2 on the $\uparrow$'s and $\downarrow$'s as long as we remember that the first position always means particle 1 and the second particle 2. **NB: The order that we write the spins in the ket now matters.** $|\uparrow\downarrow\rangle \neq |\downarrow\uparrow\rangle$. In this notation, the above relations become:

$$P_{12}|\downarrow\downarrow\rangle = |\downarrow\downarrow\rangle \tag{4.10}$$
$$P_{12}|\downarrow\uparrow\rangle = |\uparrow\downarrow\rangle \tag{4.11}$$
$$P_{12}|\uparrow\downarrow\rangle = |\downarrow\uparrow\rangle \tag{4.12}$$
$$P_{12}|\uparrow\uparrow\rangle = |\uparrow\uparrow\rangle. \tag{4.13}$$

Now, the PIP says that no matter how we label the particles, we must get the same physical description of the system. Then, if the state of the two electrons is described by $|\Psi\rangle$, it doesn't matter if we come along and act on $|\Psi\rangle$ with P_{12}, we should get a *physically equivalent* state, $|\Psi'\rangle$. Physically equivalent means that $|\Psi'\rangle$ differs from $|\Psi\rangle$ only by a *phase factor*, $e^{i\alpha}$:

$$|\Psi'\rangle = P_{12}|\Psi\rangle = e^{i\alpha}|\Psi\rangle, \tag{4.14}$$

where $\alpha \in \mathbb{R}$.

EXERCISE 4.4. Prove that states that differ only by a phase factor give rise to the same physically observable quantities. Recall that observables are calculated by taking the expectation value of Hermitian operators (though Hermiticity isn't required to solve this exercise).

EXERCISE 4.5. Prove $P_{12}P_{12} = 1$.

EXERCISE 4.6. Prove that either $\alpha = 0$ or $\alpha = \pi$.

3. Total spin or S–basis

As a consequence of the result in Prob. 50, we see that when we act on $|\Psi\rangle$ with P_{12} we must get

$$P_{12}|\Psi\rangle = \pm|\Psi\rangle, \tag{4.15}$$

so here we have the answer to the question posed earlier. If $P_{12}|\uparrow\downarrow\rangle = |\downarrow\uparrow\rangle$ and $P_{12}|\downarrow\uparrow\rangle = |\uparrow\downarrow\rangle$ then if we write (we'll describe why we label these states this way

in a moment)

$$|S_0\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) \tag{4.16}$$

$$|S_1^{(+1)}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right), \tag{4.17}$$

we get

$$P_{12}|S_0\rangle = -|S_0\rangle \tag{4.18}$$

$$P_{12}|S_1^{(0)}\rangle = |S_1^{(0)}\rangle \tag{4.19}$$

and these states on the right are physically indistinguishable from the ones on the left before application of P_{12}. The state in Eq.(4.17) is **antisymmetric** while the one in Eq.(4.20) is **symmetric**. The other states

$$|S_1^{(+1)}\rangle = |\uparrow\uparrow\rangle \tag{4.20}$$

$$|S_1^{(-1)}\rangle = |\downarrow\downarrow\rangle, \tag{4.21}$$

are already symmetric. This collection of four states, $|S_0\rangle, |S_1^{(-1)}\rangle, |S_1^{(0)}\rangle, |S_1^{(+1)}\rangle$ is called the S-basis.

We'll now set out to show through the following exercises that the states of the S–basis, as the reader may have surmised, are the states of the *total spin* of the two particles.

Exercise 4.7. Prove that the states $|S_1^{(-1)}\rangle, |S_1^{(0)}\rangle, |S_1^{(+1)}\rangle$ are eigenstates of P_{12}. What is their eigenvalue?

Exercise 4.8. Prove that the state $|S_0\rangle$ is an eigenstate of P_{12}. What is its eigenvalue?

Exercise 4.9. Calculate the normalization of all four states $\langle S_S^{(M)}|S_S^{(M)}\rangle$ using $\langle s_1' s_2'|s_1 s_2\rangle = \langle s_1'|s_1\rangle\langle s_2'|s_2\rangle$.

What is the physical meaning of the states in the S-basis? We rewrite the Hamiltonian, Eq.(4.3) as

$$\hat{H}_{spin} = -\frac{e}{m}\mathbf{S}\cdot\mathbf{B} \tag{4.22}$$

$$\mathbf{S} = \frac{1}{2}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_2), \tag{4.23}$$

where $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$ is the total (vector) spin of the two electrons.

Exercise 4.10. Show: $\mathbf{S}\cdot\mathbf{S} = \mathbf{S}^2 = \frac{3}{2} + 2\mathbf{S}_1\cdot\mathbf{S}_2$.

Next we define the **raising and lowering operators**:

$$S_{1,\pm} = \sigma_{1,\pm} = \frac{1}{2}(\sigma_{1,x} \pm i\sigma_{1,y}) = S_{1,x} \pm iS_{1,y} \tag{4.24}$$

$$S_{2,\pm} = \sigma_{2,\pm} = \frac{1}{2}(\sigma_{2,x} \pm i\sigma_{2,y}) = S_{2,x} \pm iS_{2,y} \tag{4.25}$$

$$S_{\pm} = S_{1,\pm} + S_{2,\pm}. \tag{4.26}$$

Exercise 4.11. Prove: $\mathbf{S}_1\cdot\mathbf{S}_2 = \frac{1}{2}(S_{1,+}S_{2,-} + S_{1,-}S_{2,+}) + S_{1,z}S_{2,z}$.

EXERCISE 4.12. Show that:

$$\mathbf{S}_1 \cdot \mathbf{S}_2 |S_0\rangle = -\frac{3}{4}|S_0\rangle \tag{4.27}$$

$$\mathbf{S}_1 \cdot \mathbf{S}_2 |S_1^{(-1)}\rangle = \frac{1}{4}|S_1^{(-1)}\rangle \tag{4.28}$$

$$\mathbf{S}_1 \cdot \mathbf{S}_2 |S_1^{(0)}\rangle = \frac{1}{4}|S_1^{(0)}\rangle \tag{4.29}$$

$$\mathbf{S}_1 \cdot \mathbf{S}_2 |S_1^{(+1)}\rangle = \frac{1}{4}|S_1^{(+1)}\rangle. \tag{4.30}$$

EXERCISE 4.13. Show that:

$$\mathbf{S}^2 |S_0\rangle = 0 \tag{4.31}$$

$$\mathbf{S}^2 |S_1^{(-1)}\rangle = 2|S_1^{(-1)}\rangle \tag{4.32}$$

$$\mathbf{S}^2 |S_1^{(0)}\rangle = 2|S_1^{(0)}\rangle \tag{4.33}$$

$$\mathbf{S}^2 |S_1^{(+1)}\rangle = 2|S_1^{(+1)}\rangle. \tag{4.34}$$

So we have proven that

$$\mathbf{S}^2 |S_S^{(M)}\rangle = S(S+1)|S_S^{(M)}\rangle \tag{4.35}$$

where, in the eigenvalue $S(S+1)$, S is the subscript in $|S_S^{(M)}\rangle$ and takes the values $S = 0, 1$.

EXERCISE 4.14. What operator is M in $|S_S^{(M)}\rangle$ the eigenvalue for?

EXERCISE 4.15. Calculate the eigenvalues of $\hat{H}_{spin}$ for $\mathbf{B} = (0, 0, B)$ for the states in the S-basis.

4. Addition of angular momentum

What we have been doing in the last few pages is figuring out, in a *ground-up* sort of way, how to add the angular momentum of two spin-$\frac{1}{2}$ particles in quantum mechanics. Contrast this with the situation in classical mechanics. There, we take the relevant angular momentum three-vectors and just sum them up, component by component (like we would with any three vector).

In the quantum theory, we sum up the operators of the relevant angular momenta just as we do in classical mechanics. The states are, however, a bit more involved since we can talk about, for example, each particles' individual angular momentum (the s–basis) or their total (the S–basis). Connecting these two angular momentum **bases** is, again, what we have been about these last few pages. We found that in order to find the eigenstates of the total spin in the S–basis of a pair of electrons we had to take appropriate linear combinations of the states in the s–basis. We will write them in a renamed form for consistency with the general notation. We take the states of a pair of particles in the s–basis as $|s_1 s_2; m_1 m_2\rangle$ where $s_1 = s_2 = \frac{1}{2}$ and m_1 and m_2 can be either $\pm\frac{1}{2}$. While those in the S–basis are written as $|s_1 s_2; SM\rangle$, where again $s_1 = s_2 = \frac{1}{2}$ and $S = 0, 1$ and $M = -1, 0, 1$ as we learned in the first part of this section. Sometimes when there's no chance for confusion, we'll write $|s_1 s_2; SM\rangle = |SM\rangle$. In this notation, the state of the S–basis are given in terms of the states in the s–basis as

$$|0,0\rangle = \frac{1}{\sqrt{2}}\Big(|\tfrac{1}{2}\tfrac{1}{2}; +\tfrac{1}{2} - \tfrac{1}{2}\rangle - |\tfrac{1}{2}\tfrac{1}{2}; -\tfrac{1}{2} + \tfrac{1}{2}\rangle\Big) \tag{4.36}$$

$$|1,+1\rangle = |\tfrac{1}{2}\tfrac{1}{2}; +\tfrac{1}{2} + \tfrac{1}{2}\rangle \tag{4.37}$$

$$|1,0\rangle = \frac{1}{\sqrt{2}}\Big(|\tfrac{1}{2}\tfrac{1}{2}; +\tfrac{1}{2} - \tfrac{1}{2}\rangle + |\tfrac{1}{2}\tfrac{1}{2}; -\tfrac{1}{2} + \tfrac{1}{2}\rangle\Big) \tag{4.38}$$

$$|1,-1\rangle = |\tfrac{1}{2}\tfrac{1}{2}; -\tfrac{1}{2} - \tfrac{1}{2}\rangle. \tag{4.39}$$

Recall that there are four states in the s–basis and the four states in the S–basis.

Exercise 4.16. Why?

So we can alternatively write the states in the s–basis in terms of those in the S–basis.

Exercise 4.17. Do this. Use the shorthand notation $|m_1 m_2\rangle$ for states in the s–basis and $|SM\rangle$ for states in the S–basis.

The coefficients in front of the various kets come up *all the time* and so they get a special name – angular momentum coupling coefficients – also known by the name Clebsch-Gordon coefficients, after the 19^{th} C. German mathematicians who developed them. Now, consider Fig.(1): it's a table of Clebsch-Gordon (CG) coefficients. In the upper left corner, the `1/2 × 1/2` indicates that the table is used for **coupling** $s_1 = \frac{1}{2}$ to $s_2 = \frac{1}{2}$. The body of the table consists of three sectors, delineated by thick lines – corresponding to the number (either in s–basis or S–basis) of distinct values of $S_z = s_{1,z} + s_{2,z}$.

Exercise 4.18. By looking at Fig.(1) determine the number of distinct values of S_z are there as a function of S realizing that S can take on two values: $S = 0, 1$.

Along the left-hand-side (lhs) of the body of the table, in each sector, are the values of m_1 and m_2 – the values of the individual spins of the electrons. For

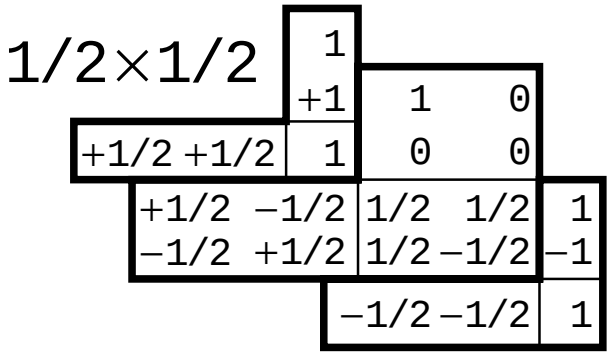

FIGURE 1. Angular momentum coupling (Clebsch-Gordon) coefficients.

example, in the first sector the lhs of the table has: `+1/2 +1/2` $\leftrightarrow |\uparrow\uparrow\rangle \leftrightarrow m_1 = +\frac{1}{2}, m_2 = +\frac{1}{2}$, and so on. Along the top side of the body of the table are the values of S (upper) and M (lower). The first sector has: ${}^{1}_{+1} \leftrightarrow |1,+1\rangle \leftrightarrow S = 1, M = +1$, and so on.

Now that we understand the layout of the table we turn to the meaning of the coefficients themselves. The first thing you need to know is that each entry in the table should be understood to have a square root symbol – *in the right place*. You should read it as:

$$\begin{aligned} \frac{1}{2} &\to +\frac{1}{\sqrt{2}} \\ -\frac{1}{2} &\to -\frac{1}{\sqrt{2}}. \end{aligned} \tag{4.40}$$

In other words, the CG coefficients are always chosen[6] to be real. To be clear: we don't take the square root of a negative entry in the table, we take minus the square root of the absolute value of the entry if it's negative.

EXERCISE 4.19. Suppose an entry of the table is $-\frac{2}{5}$. What is the CG coeff?

Now, you might already have figured out what's happening here. Suppose you want the states $|1,0\rangle$ in the S–basis. Then you go to the $M = 0$ sector (the large one in the middle of the table) and read down in the column marked ${}^{0}_{0}$. You come across a $1/2$ in the row labeled `+1/2 -1/2` and a $-1/2$ in the row labeled `-1/2 +1/2` finding, therefore that:

$$|0,0\rangle = \frac{1}{\sqrt{2}}|+\tfrac{1}{2}-\tfrac{1}{2}\rangle - \frac{1}{\sqrt{2}}|-\tfrac{1}{2}+\tfrac{1}{2}\rangle. \tag{4.41}$$

One important feature of this table, related to the fact that the **transformation** from the s–basis to the S–basis (which is what you're doing when you write the states $|SM\rangle$ in terms of the states $|s_1 s_2; m_1 m_2\rangle$) is **unitary** – we know what unitary means from before.

EXERCISE 4.20. If a matrix U is unitary, what condition does it satisfy?

[6]By appropriate selection of phase factors.

EXERCISE 4.21. Form the matrix U whose elements are the CG coefficients (rows label s–basis, columns label S–basis). Prove that this matrix is unitary.

EXERCISE 4.22. Show that the sum of the squares of the rows of the CG table are each equal to 1.

EXERCISE 4.23. Show that the sum of the squares of the columns of the CG table are each equal to 1.

So the table of CG coefficients encodes all the information about the unitary transformation from the s–basis to the S–basis. In general, if we have an orthonormal basis, like the s–basis, and we do a unitary transformation on it we get another orthonormal basis, like the S–basis.

EXERCISE 4.24. Prove that the s– and S–bases are orthormal. (Recall that a basis is orthonormal if $\langle\alpha|\alpha'\rangle = \delta_{\alpha,\alpha'}$, where α and α' label each state in the basis distinctly.)

EXERCISE 4.25. From the Clebsch-Gordon table derive the expressions for the states $|s_1 s_2; m_1 m_2\rangle$ in terms of $|SM\rangle$. Show that they are equivalent to the results for the expressions you got in Prob.(4.17)

4.1. Formal relations for two spin-$\frac{1}{2}$ particles. Now that we understand how these relations all work in detail, let's take a step back and write this information in a more formal, compact notation.

The states in the S–basis are linearly independent combinations of the states in the s–basis. This can be formally written as:

$$|s_1 s_2; SM\rangle = \sum_{m_1=-\frac{1}{2}}^{+\frac{1}{2}} \sum_{m_2=-\frac{1}{2}}^{+\frac{1}{2}} |s_1 s_2; m_1 m_2\rangle\langle s_1 s_2; m_1 m_2|s_1 s_2; SM\rangle. \tag{4.42}$$

On the lhs, we have the states, $|s_1 s_2; SM\rangle$, on the rhs we have a double sum over the states in the $|s_1 s_2; m_1 m_2\rangle$ times some quantities $\langle s_1 s_2; m_1 m_2|s_1 s_2; SM\rangle$ – which are just the CG coefficients from the table. The factor $\langle s_1 s_2; m_1 m_2|s_1 s_2; SM\rangle$ on the rhs is just an overlap of the states in the s–basis with those in the S–basis. So we could equally well write

$$|s_1 s_2; SM\rangle = \left(\sum_{m_1=-\frac{1}{2}}^{+\frac{1}{2}} \sum_{m_2=-\frac{1}{2}}^{+\frac{1}{2}} |s_1 s_2; m_1 m_2\rangle\langle s_1 s_2; m_1 m_2| \right) |s_1 s_2; SM\rangle, \tag{4.43}$$

which proves that

$$\sum_{m_1=-\frac{1}{2}}^{+\frac{1}{2}} \sum_{m_2=-\frac{1}{2}}^{+\frac{1}{2}} |s_1 s_2; m_1 m_2\rangle\langle s_1 s_2; m_1 m_2| = \mathbb{1}. \tag{4.44}$$

Be careful to note that the labels s_1 and s_2 just tell you that you're coupling spin-$\frac{1}{2}$ to spin-$\frac{1}{2}$ and they *are not* summed over. Only the m_1 and m_2 are summed over in these expressions. Also note that the sums over m_1 and m_2 only have non-zero contributions when $M = m_1 + m_2$ – this happens for the same reason that the only states appearing on the rhs of Eqs.(4.36) in the s–basis have $m_1 + m_2 = M$ on the lhs. So, for example, the coefficient $\langle \frac{1}{2}\frac{1}{2}; m_1 = +\frac{1}{2}, m_2 = +\frac{1}{2}|\frac{1}{2}\frac{1}{2}; S = 1, M = 0\rangle = 0$. You get an equivalent set of relations if you take the restrict the sums over

m_1 and m_2 subject to $m_1 + m_2 = M$. Eq.(4.44) is known as the **completeness relation**. [7]

EXERCISE 4.26. Calculate the quantities $\langle s_1 s_2; m_1 m_2 | s_1 s_2; SM\rangle \equiv \langle m_1 m_2 | SM\rangle$ for all m_1, m_2 and M which give a non-zero result by taking overlaps of states in the s–basis (as bras) with states in the S–basis (as kets).

The notation in Eq.(4.44) is very compact. It says that if you sum over the direct products of kets and bras given in the expression you obtain the unit matrix.

EXERCISE 4.27. What is the rank of the unit matrix in Eq.(4.44)?

There is, of course, a similar relation for the states of a single spin:

$$\sum_{m=-\frac{1}{2}}^{+\frac{1}{2}} |sm\rangle\langle sm| = |\tfrac{1}{2}+\tfrac{1}{2}\rangle\langle\tfrac{1}{2}+\tfrac{1}{2}| + |\tfrac{1}{2}-\tfrac{1}{2}\rangle\langle\tfrac{1}{2}-\tfrac{1}{2}|$$
$$= \mathbb{1}_{2\times 2}, \tag{4.45}$$

but note that now, the unit matrix, $\mathbb{1}$ is 2×2 or rank–2.

EXERCISE 4.28. By writing the states $|\uparrow\rangle$ and $|\downarrow\rangle$ as column vectors and the states $\langle\uparrow|$ and $\langle\downarrow|$ as a row vectors, show that Eq.(4.45) is satisfied. The direct product of a column vector $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and a row vector $v = (v_1, v_2)$ results in a matrix, M whose elements are given by $M_{ij} = u_i \times v_j$.

NB: Don't forget that $\langle\psi|$ means transpose *and complex conjugate* of $|\psi\rangle$ – but that this isn't important in the above EXERCISES since all the components are *real*.

The relations that you derived in Exs.(4.22) & (4.23) also have formal relations. These are derived by considering the orthonormality of the states in the S–basis:

$$\langle s_1 s_2; SM' | s_1 s_2; SM\rangle = \delta_{M',M}. \tag{4.46}$$

Next, we use the completeness relation Eq.(4.44) (or 'insert a complete set of s–basis states) as:

$$\langle s_1 s_2; SM' | s_1 s_2; SM\rangle = \langle s_1 s_2; SM' | \mathbb{1}_{2\times 2} | s_1 s_2; SM\rangle = \delta_{M',M} \tag{4.47}$$

$$= \langle s_1 s_2; SM' | \left[\sum_{m_1=-\frac{1}{2}}^{+\frac{1}{2}} \sum_{m_2=-\frac{1}{2}}^{+\frac{1}{2}} |s_1 s_2; m_1 m_2\rangle\langle s_1 s_2; m_1 m_2| \right] |s_1 s_2; SM\rangle = \delta_{M',M}, \tag{4.48}$$

$$= \sum_{m_1=-\frac{1}{2}}^{+\frac{1}{2}} \sum_{m_2=-\frac{1}{2}}^{+\frac{1}{2}} \langle s_1 s_2; SM' | s_1 s_2; m_1 m_2\rangle\langle s_1 s_2; m_1 m_2 | s_1 s_2; SM\rangle = \delta_{M',M}, \tag{4.49}$$

which becomes, for $M' = M$

$$\sum_{m_1=-\frac{1}{2}}^{+\frac{1}{2}} \sum_{m_2=-\frac{1}{2}}^{+\frac{1}{2}} |\langle s_1 s_2; SM | s_1 s_2; m_1 m_2\rangle|^2 = 1, \tag{4.50}$$

[7] Another example of the 'Old German Trick.'

since $\delta_{M,M} = 1$ and $\langle s_1 s_2; SM | s_1 s_2; m_1 m_2 \rangle^* = \langle s_1 s_2; m_1 m_2 | s_1 s_2; SM \rangle$.

EXERCISE 4.29. Prove this equation for $M = 0$, $S = 0, 1$ using the CG table.

This pretty much exhausts the topic of the addition of angular momentum for two spin-$\frac{1}{2}$ particles. Our project requires the wave functions of three quarks – *ie.* three spin-$\frac{1}{2}$ particles. So we need to generalize the above relations to be able to take the addition of the third particle – electron or quark (or any spin-$\frac{1}{2}$ particle) – into account. We'll do this in the next section.

Let's review. We began by considering two spin-$\frac{1}{2}$ particles and "coupling" ("adding") their spins together. This yielded the singlet with total spin $S = 0$ and the triplet with total spin $S = 1$. In shorthand notation:

$$\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1 \tag{4.51}$$

or

$$\mathbf{2} \otimes \mathbf{2} = \mathbf{1} \oplus \mathbf{3} \tag{4.52}$$

where in the first equation we have written the coupling in terms of the total spin and the second we have written the coupling in terms of the *dimensionality* of the bases, **2** on the lhs for the s–basis and **1** for the singlet and **3** on the rhs for the S–basis. Eq.(4.51) is non-arithmetic – treating the spins as real numbers and using $\otimes$ and $\oplus$ as ordinary multiplication and addition doesn't yield a correct equation. But the second equation, Eq.(4.52) does. ("Two times two equals 1 plus 3.") And this is a good way to check your calculations. [8]

Note further that the bold numbers in Eq.(4.52) are simply 2×(# in Eq.(4.51))+ 1. For example, for $S = \frac{1}{2}$ there are $2S + 1 = 2$ states with different values of M. In this way, you can check that after you couple together two angular momenta, say j_1 and j_2 you get the right number of states of total angular momentum.

Let's see how this works by generalizing our results for the coupling of two spin-$\frac{1}{2}$ particles to the coupling of two arbitrary angular momenta. By arbitrary we mean either angular momentum j_1 and/or j_2 can be integral (0,1,2,...) or half odd-integral ($\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$). The allowable total angular momenta, which we state without proof (since it requires a bit more group theory than we have time for) is:

$$j_1 \otimes j_2 = (j_1 + j_2) \oplus (j_1 + j_2 - 1) \oplus (j_1 + j_2 - 2) \oplus \cdots \oplus |j_1 - j_2|, \tag{4.53}$$

where the absolute value $|j_1 - j_2|$ is needed because the magnitude of angular momentum is always a positive number and we're allowing either $j_1 \leq j_2$ or $j_1 \geq j_2$. Let's refer to this relation as the 'vector couplings.'

EXERCISE 4.30. Setting $j_1 = j_2 = \frac{1}{2}$ show that this reduces to Eq.(4.51).

EXERCISE 4.31. In the table below, work out the vector couplings for the (j_1, j_2) in the leftmost column and mark the column with an 'X' in whenever these couple to the total j given in the top line of the table. The first two rows have been filled in as an example.

[8] Unfortunately, as in many cases in physics' notation, there is no universal notation in this case and you're likely to see both.

(j_1, j_2)	0	$\frac{1}{2}$	1	$\frac{3}{2}$	2	$\frac{5}{2}$	3	$\frac{7}{2}$	4
$(0,\frac{1}{2})$		X							
$(\frac{1}{2},\frac{1}{2})$	X		X						
$(0,1)$									
$(\frac{1}{2},1)$									
$(1,1)$									
$(0,2)$									
$(\frac{1}{2},2)$									
$(1,2)$									
$(\frac{3}{2},2)$									
$(2,2)$									

Notice that when we couple j_1 and j_2 the resultant j is always either integral or half odd-integral. Also notice that when we couple integral with integral the resultant is integral: $1 \otimes 1 = 0 \oplus 1 \oplus 2$. And that when we couple integral to half odd-integral we obtain half odd-integral. Etc.

EXERCISE 4.32. Couple $j_1 = 1$ to $j_2 = \frac{1}{2}$ to get $|jm\rangle$. Work out all the states – there are six – using the CG table I sent along with these notes in the file `cgcoeffs.pdf`. Use:

$$|j_1 j_2; jm\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1 j_2; m_1 m_2\rangle \langle j_1 j_2; m_1 m_2 | j_1 j_2; jm\rangle \tag{4.54}$$

with $j_1 = 1$, $j_2 = \frac{1}{2}$ and $j = \frac{3}{2}, \frac{1}{2}$.

5. Three spin-$\frac{1}{2}$ particles

The last problem had you derive the 6 states with 'good' total angular momentum $j = \frac{1}{2}, \frac{3}{2}$. By 'good' we mean "having a definite value" – which is another way of saying the state is an *eigenstate* of the total angular momentum operator.

Let's list these states. We use the notation $|j_1 j_2 j_2, jm; S_3\rangle = |jm; S_3\rangle$, where now we have an additional label, S_3 to further distinguish the states of good total angular momentum. This is needed because we get *two* distinct $j = \frac{1}{2}$ sets of states. One from coupling $j_1 = 0$ with $j_2 = \frac{1}{2}$ and one from $j_1 = 1$ with $j_2 = \frac{1}{2}$. Let's use the results from the last problem to write down the spin states and for three spin-$\frac{1}{2}$

particles then discuss the notation.

$$|\tfrac{3}{2},+\tfrac{3}{2};S\rangle = |\uparrow\uparrow\uparrow\rangle \tag{4.55}$$

$$|\tfrac{3}{2},+\tfrac{1}{2};S\rangle = \frac{1}{\sqrt{3}}|\uparrow\uparrow\downarrow\rangle + \frac{1}{\sqrt{3}}|\uparrow\downarrow\uparrow\rangle + \frac{1}{\sqrt{3}}|\downarrow\uparrow\uparrow\rangle \tag{4.56}$$

$$|\tfrac{3}{2},-\tfrac{1}{2};S\rangle = \frac{1}{\sqrt{3}}|\downarrow\downarrow\uparrow\rangle + \frac{1}{\sqrt{3}}|\downarrow\uparrow\downarrow\rangle + \frac{1}{\sqrt{3}}|\uparrow\downarrow\downarrow\rangle \tag{4.57}$$

$$|\tfrac{3}{2},-\tfrac{3}{2};S\rangle = |\downarrow\downarrow\downarrow\rangle \tag{4.58}$$

$$\begin{aligned}|\tfrac{1}{2},+\tfrac{1}{2};\rho\rangle &= |0,0\rangle \otimes |\uparrow\rangle \\ &= \frac{1}{\sqrt{2}}|\uparrow\downarrow\uparrow\rangle - \frac{1}{\sqrt{2}}|\downarrow\uparrow\uparrow\rangle\end{aligned} \tag{4.59}$$

$$\begin{aligned}|\tfrac{1}{2},-\tfrac{1}{2};\rho\rangle &= |0,0\rangle \otimes |\downarrow\rangle \\ &= \frac{1}{\sqrt{2}}|\uparrow\downarrow\downarrow\rangle - \frac{1}{\sqrt{2}}|\downarrow\uparrow\downarrow\rangle\end{aligned} \tag{4.60}$$

$$\begin{aligned}|\tfrac{1}{2},+\tfrac{1}{2};\lambda\rangle &= \sqrt{\frac{2}{3}}|1,+1\rangle \otimes |\downarrow\rangle - \frac{1}{\sqrt{3}}|1,0\rangle \otimes |\uparrow\rangle \\ &= \frac{2}{\sqrt{6}}|\uparrow\uparrow\downarrow\rangle - \frac{1}{\sqrt{6}}|\uparrow\downarrow\uparrow\rangle - \frac{1}{\sqrt{6}}|\downarrow\uparrow\uparrow\rangle\end{aligned} \tag{4.61}$$

$$\begin{aligned}|\tfrac{1}{2},-\tfrac{1}{2};\lambda\rangle &= \frac{1}{\sqrt{3}}|1,0\rangle \otimes |\downarrow\rangle - \sqrt{\frac{2}{3}}|1,-1\rangle \otimes |\uparrow\rangle \\ &= \frac{1}{\sqrt{3}}|\uparrow\downarrow\downarrow\rangle + \frac{1}{\sqrt{3}}|\downarrow\uparrow\downarrow\rangle - \frac{2}{\sqrt{6}}|\downarrow\downarrow\uparrow\rangle.\end{aligned} \tag{4.62}$$

The label which we've called S_3 takes on three distinct values, S, ρ, and λ. They correspond to the characteristics of the states under **permutations**. We've learned about one type of permutation – the transposition operator, P_{ij} which interchanges the labels of particles i and j. We're not going to get into a full discussion of the **permutation group on n objects**, S_n here. We'll just note the following: states with $S_3 = S$ are completely symmetric – we get the same state if we make *any* permutation of the particle labels. For example if we apply P_{23} to $|\frac{3}{2},+\frac{1}{2};S\rangle$ we obtain

$$\begin{aligned}P_{23}|\tfrac{3}{2},+\tfrac{1}{2}\rangle &= \frac{1}{\sqrt{3}}|\uparrow\downarrow\uparrow\rangle + \frac{1}{\sqrt{3}}|\uparrow\uparrow\downarrow\rangle + \frac{1}{\sqrt{3}}|\downarrow\uparrow\uparrow\rangle \\ &= \frac{1}{\sqrt{3}}|\uparrow\uparrow\downarrow\rangle + \frac{1}{\sqrt{3}}|\uparrow\downarrow\uparrow\rangle + \frac{1}{\sqrt{3}}|\downarrow\uparrow\uparrow\rangle \\ &= |\tfrac{3}{2},+\tfrac{1}{2}\rangle.\end{aligned} \tag{4.63}$$

These states are relatively simple concerning their properties under permutations of the particle labels. The $j = \frac{1}{2}$ states however are more complicated. For now, we just note that since they were built from the S–basis states, which are antisymmetric under the action of P_{12} for $S = 0$ and symmetric under P_{12} for $S = 1$, the states $|jm;S_3\rangle$ inheret these properties. The labels ρ and λ indicate this inhereted symmetry property. The state ρ is antisymmetric under the action of P_{12} since it was built from the $S = 0$ two-particle state and the state λ is symmetric under the action of P_{12} since it was built from the $S = 1$ two particle state. Note that the action of the other permutations, for example, P_{23} are *not simple* when acting on

the $S = \rho, \lambda$ states. Though the results of applying *any* permutation operator are linear combinations of the the basis states.[9]

We'll need to be aware of the symmetry properties of the states when we come to combine these spin states with states that describe the other quantum numbers of the baryons, like **isospin**. For now, just keep in mind that like the two-particle states in the S–basis we can characterize the three-particle spin states by their symmetry properties under permutations.

Let's study the spin structure of these states in more detail in terms of their angular momentum. The analog of Eq.(4.35) and the action of the operator J_z on the above states can be written

$$\mathbf{J}^2|jm\rangle = j(j+1)|jm\rangle \tag{4.64}$$

$$J_z|jm\rangle = m|jm\rangle \tag{4.65}$$

where $\mathbf{J}\cdot\mathbf{J} \equiv \mathbf{J}^2$. Be sure to note that here

$$\mathbf{J}^2 \equiv J_x^2 + J_y^2 + J_z^2 \tag{4.66}$$

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 + \mathbf{J}_3 \tag{4.67}$$

Equations (4.64) and (4.65) mean that the state $|jm\rangle$ are *eigenstates* of the operators $\mathbf{J}^2$ and J_z, respectively. And so they have good values of total angular momentum.

EXERCISE 4.33. Are the states $|jm\rangle$ eigenstates of $\mathbf{J}_1^2$? What about of $J_{1,z}$?

In fact, while the basis $|j_1 j_2 j_3; m_1 m_2 m_3\rangle$ (the analog of the two-particle s–basis) are eigenstates of the all the operators $\mathbf{J}_1^2, \mathbf{J}_2^2, \mathbf{J}_3^2, J_{1,z}, J_{2,z}, J_{3,z}$, the states $|j_1 j_2 j_3; jm\rangle$ (the analog of the two-particle S–basis states) are eigenstates of the operators $\mathbf{J}_1^2, \mathbf{J}_2^2, \mathbf{J}_3^2, \mathbf{J}^2, J_z$. These two sets of operators are referred to as **complete sets of commuting observables**. They afford a complete physical description of the system under study.

That's about all we'll say about the three-particle spin basis for now. There is a lot more one can say about this subject – it's vast. For example, you can define operators, $J_\pm$ like $S_\pm$ of Eq.(4.24) which have the effect of raising and lowering m in $|jm\rangle$ as follows

$$J_\pm|j,m\rangle = \sqrt{(j\mp m)(j\pm m+1)}|j,m\pm 1\rangle, \tag{4.68}$$

where $J_\pm = J_x \pm iJ_y$ and remember that here, for example, $J_x = J_{1,x} + J_{2,x} + J_{3,x}$. So you can see that if you know one of the states in the total J basis you can find the other by repeated application of the operators $J_\pm$.

EXERCISE 4.34. Apply the operator J_- to the **stretched state** $j = m$, ie. $|j = \frac{3}{2}, m = j = +\frac{3}{2}\rangle$ to obtain the state $|\frac{3}{2}, +\frac{1}{2}\rangle$. You need to apply $J_- = J_{1,-} + J_{2,-} + J_{3,-}$, using J_- on the lhs and $J_{1,-} + J_{2,-} + J_{3,-}$ on the rhs.

Now that we have the spin states of three quarks somewhat under control, we need to turn to the other quantum numbers of the quarks – **flavor** and **color**.

---

[9]For the ambitious reader, this line of argument leads directly to matrix representations of the the symmetric group, S_n when acting on a basis of states. See if you can, for example, find the matrices corresponding to P_{12}. P_{13}, and P_{23} from the states in Eqs.(4.59) & (4.61). Do these matrices satisfy the S_3 group multiplication rules, for example: $P_{12}P_{13} = P_{23}P_{12}$?

6. The full three quark wave function

6.1. Single quark revisited. The quark has three quantum numbers which describe its physical properties: *spin, flavor, and color.*[10] We've covered the spin, $s = \frac{1}{2}$. The other two quantum numbers have been described in your readings in Griffiths and other places. They are somewhat more abstract, physically speaking, than the spin, since most people have an intuitive sense of what spin "is". (But remember that the quark spin doesn't arise because the electron is really physically spinning like the earth on its axis.) This is because the spin of the quark is "in" physical space. You can 'see' it. Flavor and color on the other hand are not. They are therefore often referred to as **internal quantum numbers**. But beware of this moniker. *They are no less real than spin.* One nice thing about them is that, mathematically, they aren't too different from spin. The internal quantum numbers describe other physical attributes of the quarks which have been inferred on the basis of *experimental observation* – they must be there, and be described in just the right mathematical fashion, to describe what is observed in nature.

Let's denote the single quark state as $|q\rangle = |\sigma, \tau, c\rangle$, q stands collectively for σ, τ, c which are the **quark coordinates**. Here σ is the spin-projection, what we've been calling m or m_1 or M, etc. $(\uparrow, \downarrow)$ for the single quark. The label τ (rhymes with 'Ow!') is for the flavor $(u, d, \ldots)$ and the label c describes the color (below) of the quark. There are 6 possible flavors of quarks that we know of. This means τ can have the 'values': u, d, s, c, b, t often referred to by the names up, down, strange, charm, bottom, top (Brits prefer beauty and truth for the last two). And each flavor of quark comes in three colors, $c = R, G, B$ (sometimes these are R, Y, B – it doesn't matter). These labels, of course, aren't real flavors or colors. They're just names that people have come up with to describe the observed possible quantum states that quarks can occupy. Just like the spin of a single quark can be only $\pm\frac{1}{2}$ the flavor of the quark can be only u, d, s, c, b, t and the color can be only R, G, B. Now, for the purposes of our study of hadronic structure, we won't be interested in the c, b, or t quarks – they're too heavy to play much of a role in the light or 'low-lying' hadrons that we want to study.

Exercise 4.35. Look up the masses of the all the quarks and tabulate them, noting their flavor, charge, and mass. Plot their masses on a semi-logarithmic graph where the abscissa (y axis) is the logarithm of the mass and the ordinate is the flavor. Calculate the mass of the top quark in grams.

So how do we describe these quarks mathematically? Well, think back to the case of the the spin of a single quark. We described it in terms of a single quantum number, m (or now σ) which we represented mathematically as a particular of a spinor – a complex, 2-D vector. Now, each quark is going to be described in terms of the three quantum numbers, the possible values of which span a $2 \times 3 \times 3 = 18$ dimensional *complex* space. We multiply the dimensions because the quark is simultaneously in a state of a given spin, flavor and color. So the generalization of

[10] Remember that in the quantum theory, all particles of a given species are identical. So we speak of *the* quark.

the spinors that we used to describe the spin of the quark look like

$$|q\rangle = \begin{bmatrix} q_{\uparrow uR} \\ q_{\downarrow uR} \\ q_{\uparrow dR} \\ \vdots \\ q_{\uparrow uG} \\ \vdots \\ q_{\downarrow sB} \end{bmatrix}, \tag{4.69}$$

$$= \sum_{\sigma=\uparrow,\downarrow} \sum_{\tau=u,d,s} \sum_{c=R,G,B} q_{\sigma,\tau,c} |\sigma, \tau c\rangle \tag{4.70}$$

where we haven't shown all the entries in the 18-dimesional complex vector $|q\rangle$ required to describe the quark. Note that each entry, $q_{\sigma,\tau,c} \in \mathbb{C}$ – *ie.* is a complex number. And that the probability of finding a quark with the particular values of σ, τ, c are given by the norm (or *modulus*) of this entry

$$\mathcal{P}_{\sigma,\tau,c} = |\langle \sigma, \tau, c|q\rangle|^2 \tag{4.71}$$

$$= |q_{\sigma,\tau,c}|^2 \tag{4.72}$$

just like in the case of the spin – compare to Eqs.(1.102).

Phew! That thing is crazy! Look at the size of it. Especially when you consider that you're going to be working with not one by *three* quarks with $18^3 = 5832$ components! You might be terrified by such an object. You might think of hanging up your pencil and forgetting about this project. But have no fear. Things are not all that bad. In practice, we never have to deal with such and unwieldy object. We always deal with only one sector – spin, flavor or color, at a time. This and some other tricks will simplify our task.

6.2. Flavor states. In the last section, we decided to confine our attention to the light quarks, u, d, and s. Let's make it even more simple. Let's talk only about u and d quarks.

Now, if u quark and a d quark have *exactly* the same mass: $m_u = m_d$ then since they all their other physical attributes (there's only spin!) are the same we should properly, in the quantum theory, treat them simply as different *states* of the same particle. This is similar to what happens in the case of the intrinsic spin. We don't speak of a spin $\uparrow$ electron being a different particle from a spin $\downarrow$ quark! That even sounds ridiculous. In this way we associate the two different flavors u and d with different *flavor states*

$$|u\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad |d\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{4.73}$$

But these are just exactly the same as the basis states we used to describe the spin of the quark, with the replacement $\uparrow \rightarrow u$ and $\downarrow \rightarrow d$. So:

> The mathematics of flavor symmetries are identical to those of rotation symmetries, namely $SU(2)$.

This time, however, rotations of the quark flavor don't occur in the physical configuration space, $\mathbb{R}^3$, they occur in flavor space, which is a Hilbert space which

we'll call $\mathcal{F}$. This space is called *isospin space*[11] and everything gets generalized from 'spin' to 'isospin' – $S \to T$, $S_z \to T_z$, $\mathbf{S}^2 \to \mathbf{T}^2$, ... For example, T is the total isospin, etc.

EXERCISE 4.36. For three quarks, write down the $T = \frac{1}{2}$ isospin states for both the ρ and λ sets. Remember $\uparrow \to u$ and $\downarrow \to d$!

Easy, since we've done all the hard work for spin.

6.3. Color states. The last quantum number we need to contend with is *color*. Color comes in three varieties and since you've probably figured out the game by now, you could guess that the three states corresponding to color can be taken to be:

$$|R\rangle = \begin{bmatrix}1\\0\\0\end{bmatrix}, \qquad |G\rangle = \begin{bmatrix}0\\1\\0\end{bmatrix}, \qquad |B\rangle = \begin{bmatrix}0\\0\\1\end{bmatrix} \tag{4.74}$$

and the general color state of a single quark is

$$|\chi^C\rangle = \sum_{c=R,G,B} \chi^C_c |c\rangle, \tag{4.75}$$

where $\chi^C_c \in \mathbb{C}$. Note that since there are *three* basis states, $|R\rangle, |B\rangle, |G\rangle$ the group theory that describes color is distinct from that which describes spin or isospin. Spin and isospin were two-dimensional and we used $SU(2)$ to describe it.

EXERCISE 4.37. What is the group which describe u, d, s quarks (assuming they all have the same mass)?

When we consider color states of more than one quark, we can still classify them by their symmetry properties. One out of the many three quark color states one is special.

EXERCISE 4.38. How many three quark color states are there?

It is completely antisymmetric under any exchange of the particle labels. It is called the **color singlet** and it is given by

$$|\chi^C_0\rangle = \frac{1}{\sqrt{6}}\Big(|RGB\rangle - |RBG\rangle + |GBR\rangle - |BGR\rangle + |BRG\rangle - |GRB\rangle\Big), \tag{4.76}$$

$$= \frac{1}{\sqrt{6}}\epsilon_{\alpha\beta\gamma}|c^\alpha_1 c^\beta_2 c^\gamma_3\rangle, \tag{4.77}$$

where the subscript, '0' indicates that this state is the 'singlet.' (The indices α, β, γ each run over $R, G, B \leftrightarrow 1, 2, 3$ and they are summed over.)

EXERCISE 4.39. When we rotate the spin singlet state of two quarks [Eq.(4.36)] nothing happens to it. Why? What happens to the color singlet state when we 'rotate' it in color space?

EXERCISE 4.40. Calculate $\langle\chi^C_0|\chi^C_0\rangle$.

EXERCISE 4.41. Show that $P_{ij}|\chi^C_0\rangle = -|\chi^C_0\rangle$ for $ij = 12, 13, 23$.

[11] When we have just two flavors u and d we speak of 'isospin space'. With three flavors, u, d, s we speak of 'flavor space'. But u and d can be referred to as 'flavor' in either case.

What's so special about the color singlet? All baryons (that is, all states of three quarks) that have been observed to date are apparently, to very high accuracy, color singlets. This is an astonishing fact. It means that all the baryons (in fact, all the hadrons) are equal parts R, G, and B. This means that all physical states are *colorless*. This is what it means to be 'color singlet.' No matter what transformations you do in color space (analogous to rotations in physical space and flavor space) the effect on a color singlet state is *nil* – nothing happens to a color singlet. For our purposes, it means that we can write down the state once and *forget about it*! That's right, we don't even explicitly write down the color singlet state when we're writing down the wave function. We just remember that it's there in case we need it. We won't.

6.4. Spin-isospin-color wave function. Now we come into the home stretch. We need to write down the complete wave function for states of three quarks for the baryon, $|B\rangle$ where B is the baryon label: $B = N, \Delta$ (we'll consider $\Sigma, \Xi, \Lambda, \Omega$ later). (N means p or n.) Each baryon is described by a state which is a sum of products of spin, isospin, and color factors

$$|B; S, S_z, T, T_z, C = 0\rangle = \left(\sum_{S_3^{(S)}, S_3^{(T)}} C^{(B)}_{S_3^{(S)}, S_3^{(T)}} |\chi^S(S, S_z; S_3^{(S)})\rangle \otimes |\chi^T(T, T_z; S_3^{(T)})\rangle \right) \otimes |\chi_0^C\rangle. \tag{4.78}$$

Here, $C = 0$ on the lhs means that we're considering the color singlet state. The color singlet *factors* from the rest of the wave function. So we can write:

$$|B; S, S_z, T, T_z, C = 0\rangle = |B; S, S_z, T, T_z\rangle \otimes |\chi_0^C\rangle, \tag{4.79}$$

$$|B; S, S_z, T, T_z\rangle = \sum_{S_3^{(S)}, S_3^{(T)}} C^{(B)}_{S_3^{(S)}, S_3^{(T)}} |\chi^S(S, S_z; S_3)\rangle \otimes |\chi^T(T, T_z; S_3)\rangle \tag{4.80}$$

Before we discuss this formula in detail, we recall that states composed of an odd number of fermions must be completely antisymmetric under permutation of the particle labels due to PIP. So we must take sums of the form in Eq.(4.78) in order to achieve this. As we noted earlier, the states in Eqs.(4.55) – (4.62) have complicated properties under permutations for some of the transposition operators. Then, with this in mind, for a given spin and isospin of the baryon B: S, S_z, T, T_z we take sums of different symmetry-type wave functions to obtain overall antisymmetric wave functions.

Now, we have already noted that the color singlet wave function is completely antisymmetric. Then we must have the spin-isospin factor, Eq.(4.80) be completely symmetric under particle relabelings.

Exercise 4.42. Prove that the spin-isospin factor, $\chi^S \chi^T$ is completely symmetric in Eq.(4.80).

Next, consider the sums in Eq.(4.80). These are sums over the symmetry types in Eqs.(4.55)–(4.62). We take linear combinations of these states with the coefficients $C^{(B)}_{S_3^{(S)}, S_3^{(T)}}$. These coefficients are like the CG coefficients in the sense that they relate states in one basis to states in another. In this case however the

bases that are being related are those with certain symmetry types. The indices $S_3^{(S)}$ and $S_3^{(T)}$ run over the labels S, ρ, λ. The coefficients have the values

$$C^{(N)}_{\rho,\rho} = \frac{1}{\sqrt{2}} \tag{4.81}$$

$$C^{(N)}_{\lambda,\lambda} = \frac{1}{\sqrt{2}} \tag{4.82}$$

when $B = N$ and

$$C^{(\Delta)}_{S,S} = 1, \tag{4.83}$$

when $B = \Delta$, with all the rest being zero. You have to be careful that you choose the states $|S, S_z; S_3^{(S)}\rangle$ and $|T, T_z; S_3^{(T)}\rangle$ correctly for the baryon state $|B; S, S_z, T, T_z\rangle$ that you're considering. Just follow the notation and things should work out all right.

Exercise 4.43. Calculate the overall wave function for the following states
$|N; S = \frac{1}{2}, S_z = +\frac{1}{2}, T = \frac{1}{2}, T_z = +\frac{1}{2}\rangle$
$|N; S = \frac{1}{2}, S_z = +\frac{1}{2}, T = \frac{1}{2}, T_z = -\frac{1}{2}\rangle$
$|\Delta; S = \frac{3}{2}, S_z = +\frac{3}{2}, T = \frac{3}{2}, T_z = +\frac{3}{2}\rangle$
$|\Delta; S = \frac{3}{2}, S_z = +\frac{3}{2}, T = \frac{3}{2}, T_z = -\frac{1}{2}\rangle$.

7. Magnetic moments of the baryons

7.1. Discussion. We learned in the Chapter 3 that a spin-$\frac{1}{2}$ particle in a magnetic field has an energy which depends on the orientation of the spin of the particle with respect to the field. In fact, any particle with non-zero spin has such an energy which for subatomic objects is generally a very small amount of energy. Nevertheless, experiments are constructed which permit the measurement of this small energy – or rather, differences in energies between different orientations of the spins of the particle being studied.

We already know that the energy of two independent spin-$\frac{1}{2}$particles is given by Eq.(4.3). The generalization to three particles is obvious

$$\hat{H}_{spin} = \hat{H}_{spin,1} + \hat{H}_{spin,2} + \hat{H}_{spin,3} \tag{4.84}$$

$$= -(\mu_1 \mathbf{S}_1 \otimes \mathbb{1}_2 \otimes \mathbb{1}_3 + \mu_2 \mathbb{1}_1 \otimes \mathbf{S}_2 \otimes \mathbb{1}_3 + \mu_3 \mathbb{1}_1 \otimes \mathbb{1}_2 \otimes \mathbf{S}_3) \cdot \mathbf{B}, \tag{4.85}$$

where we have made explicit the identity matrices for the "inactive" particles in each term. We won't write them from here on out. Simply put:

$$\hat{H}_{spin} = -(\mu_1 \mathbf{S}_1 + \mu_2 \mathbf{S}_2 + \mu_3 \mathbf{S}_3) \cdot \mathbf{B}, \tag{4.86}$$

The magnetic moment of a quark, μ_i is defined as

$$\mu_i = g_q \frac{q_i e}{2m_i} \tag{4.87}$$

where q_i can be positive or negative and is the charge of quark i in units of $e = |e|$ (the charge of the proton), $g_q = 2$ is the quark's "gyromagnetic ratio," and m_i is its mass. Clearly then, if the magnetic field is $\mathbf{B} = (0, 0, B)$ can write the energy of the baryon as

$$\hat{H}_{spin} = -\hat{\mu}_B \, B, \tag{4.88}$$

$$\hat{\mu}_B = \sum_{i=1}^{3} \mu_i S_{i,z}, \tag{4.89}$$

$$= \sum_{i=1}^{3} \frac{q_i e}{2m_i} \sigma_{i,z}. \tag{4.90}$$

Here we have written $\hat{\mu}_B$ with a 'carat' to show that it is an operator.

Exercise 4.44. Show that this is equivalent to Griffiths Eq.(5.117) if we take $\mu_B = \langle B \uparrow | \hat{\mu}_B | B \uparrow \rangle$.

Let's recall the physical properties of the u, d and s quarks:

$$m_u = 313 \text{ MeV} \qquad m_d = 313 \text{ MeV} \qquad m_s = 550 \text{ MeV} \tag{4.91}$$

$$q_u = +\frac{2}{3} \qquad q_d = -\frac{1}{3} \qquad q_s = -\frac{1}{3}. \tag{4.92}$$

These are the *constituent quark* of *effective* masses of the quarks. Compare with the discussion in §4.5 of Griffiths and his Table 4.4 there. We'll simply take these as 'experimental facts' at this stage and not worry about where these numbers come from.[12]

[12]In fact, the constituent masses are *not experimentally accessible* (they're not obervables) because of a long distance property of QCD called **confinement**.

The masses we're for the quarks using were originally determined from an analysis which fit the magnetic moments and masses of the baryons by taking the masses of u, d, s to be parameters. So what we're doing is a consistency check of this analysis. We: i) assume values of the quark masses; ii) check to see if the magnetic moments and the masses we obtain make sense. If this isn't completely clear at this stage, don't worry – it'll become so once you see how we calculate the magnetic moments and masses of the baryons.

So remember: the quarks each have a magnetic moment and they *combine* to give the magnetic moment of the baryon. It's all a matter of determining this combination or *coupling* of the spins and, therefore flavors (since the spin-isospin wave function must be overall symmetric) in the baryon wave function that will give us its total magnetic moment. This is familiar from the calculation of the energy of two electrons in a magnetic field. There we saw that depending on how we *combined* the spins either to give $S = 0, S_z = 0$ or $S = 1, S_z = +1$ we obtained different energies in the magnetic field.

You might bw wondering why I refer to only these two states of $S = 0$ and $S = 1$. First, by definition, the magnetic moment of the state refers to the magnetic moment calculated in the *stretched state – ie.* if we have a particle of spin J then the magnetic moment is determined by calculating the energy Eq.(4.86) in the stretched state $|JJ\rangle$. The states of different $M \neq J$ are related to the stretched state $|JJ\rangle$ by a rotation of *all three quarks* or, equivalently of the magnetic field while the quarks are held still. These states, therefore, don't tell us anything new about the *intrinsic* structure of the baryon and that's what we're interested in.

7.2. Magnetic moment evaluation. This subsection will deal with the details of calculating the magnetic moment operator for a proton in the state $|p\uparrow\rangle$. A *one-body operator* $\mathcal{O}^{(1)}$, like the magnetic moment operator $\hat{\mu}_B$, is a sum of terms each of which acts on the coordinates of just one particle. Contrast this with operators like

$$\mathcal{O}^{(2)} = \sum_{i<j=1}^{3} V_0 \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \tag{4.93}$$

$$= V_0(\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 + \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_3 + \boldsymbol{\sigma}_2 \cdot \boldsymbol{\sigma}_3) \tag{4.94}$$

in which each term has operators which act the coordinates of a pair of particles. We'll come back to this sort of operator in the next subsection when we discuss the baryon mass operator.

Returning to the baryon magnetic moment operator, we have

$$\hat{\mu}_B = \sum_{i=1}^{3} \frac{q_i e}{2m_i} \sigma_{i,z}. \tag{4.95}$$

In Ex.(4.43) you found the wave function of the spin $\uparrow$ proton to be:

$$\begin{aligned} |p\uparrow\rangle = \frac{1}{3\sqrt{2}}\Big[&2|u\uparrow u\uparrow d\downarrow\rangle - |u\uparrow u\downarrow d\uparrow\rangle - |u\downarrow u\uparrow d\uparrow\rangle \\ &+2|u\uparrow d\downarrow u\uparrow\rangle - |u\downarrow d\uparrow u\uparrow\rangle - |u\uparrow d\uparrow u\downarrow\rangle \\ &+2|d\downarrow u\uparrow u\uparrow\rangle - |d\uparrow u\uparrow u\downarrow\rangle - |d\uparrow u\downarrow u\uparrow\rangle\Big]. \end{aligned} \tag{4.96}$$

Let's talk about this equation in some detail to make sure we're properly oriented.

In this equation, we are writing the quantum numbers of quark '1' first, then those of quark '2', etc. You should think of the flavor and spin quantum numbers for a given quark as a 'unit' like $|\tau\sigma\rangle$. For example, the first term is more completely written

$$\frac{2}{\sqrt{18}}|(u\uparrow)_1(u\uparrow)_2(d\downarrow)_3\rangle. \tag{4.97}$$

You can see that writing the quantum numbers of the quarks in the order of their labelings saves a lot of writing.

The wave function can be written in a form which is a little easier on the eyes:

$$|p\uparrow\rangle = \frac{1}{3\sqrt{2}}\Big[2|u\uparrow u\uparrow d\downarrow\rangle - |u\uparrow u\downarrow d\uparrow\rangle - |u\downarrow u\uparrow d\uparrow\rangle \\ +\text{cyclic permutations}\Big]. \tag{4.98}$$

Here we have dropped the second and third lines of Eq.(4.96) and instead written 'cyclic permutations'. Cyclic permutations are those permutations which change all of the indices simultaneously.

EXERCISE 4.45. Prove that for S_3 the number of independent cyclic permutations is two.[13] Give them as follows: if we start with the ordering 123, give the resultant orderings of the two possible cyclic permutations.

EXERCISE 4.46. **Extra credit** Prove that the number of cyclic permutations in S_n is $(n-1)!$.

We're now in a position to calculate. The magnetic moment of the proton μ_p is defined as the expectation value of the operator in Eq.(4.88) in the stretched state of the proton.

EXERCISE 4.47. The proton has $S = \frac{1}{2}$. What is the M_S of the stretched state?

Then μ_p is given by the expectation value

$$\mu_B = \langle\hat{\mu}_B\rangle = \langle p\uparrow|\hat{\mu}_B|p\uparrow\rangle \tag{4.99}$$

$$= \langle p\uparrow|\sum_{i=1}^{3}\frac{q_i e}{2m_i}\sigma_{i,z}|p\uparrow\rangle. \tag{4.100}$$

Well, this looks like a big job. We have three terms in the operator $\mathcal{O}^{(1)}$, one for each particle. And each of these terms requires, ostensibly[14], the calculation of 81 matrix elements[15] since there are 9 terms in the bra and 9 terms in the ket. Let's see if we can't simplify the task somewhat.

The first thing we notice is that the wave function $|p\uparrow\rangle$ for the proton is a sum of 9 terms each of which is *orthogonal* to the other 8 terms.

EXERCISE 4.48. Prove that the off-diagonal matrix elements, like $\langle u\uparrow u\uparrow d\downarrow|\sigma_{i,z}|u\downarrow u\uparrow d\uparrow\rangle$ is zero no matter what value i (particle number) takes.

EXERCISE 4.49. By eliminating the off-diagonal matrix elements how many terms do we have to calculate now?

[13]Remember that the notation S_n means the symmetric group of n objects.

[14]Meaning: 'not taking into account possible orthogonality of inactive particles.'

[15]Here "matrix elements" is used to mean either 'bra different from ket' or 'bra same as ket.'

So we've considerably simplified our task and learned the first analytical tool for evaluating operators on many-body wave functions: always determine which matrix elements don't contribute due to orthogonality of the *inactive* coordinates – inactive coordinates are those of the particles which are not being acted upon by the operator under consideration.

We can further simplify the calculation however, by exploiting the redundancy in the proton wave function's cyclic pertmutations in Eq.(4.98). We can write the cyclic permutations as operators. In the case of the transposition operators we wrote P_{12} (described in the text near Eq.(4.6)) to mean $1 \to 2, 2 \to 1$. For the cyclic operators in S_3, we want $1 \to 2, 2 \to 3, 3 \to 1$ and $1 \to 3, 3 \to 2, 2 \to 1$. The first is written as P_{123} while the second is P_{132}. You should read them as "1 to 2 to 3 to 1" and "1 to 3 to 2 to 1". Note that, as advertised, none of the indices remain the same under cyclic permutation.

EXERCISE 4.50. List the cyclic permutations for S_4.

Using these operators we can rewrite proton wave function as

$$|p\uparrow\rangle = [1 + P_{123} + P_{132}]\,|\Psi_{uud}\rangle \tag{4.101}$$

$$|\Psi_{uud}\rangle = \frac{1}{3\sqrt{2}}\left[2|u\uparrow u\uparrow d\downarrow\rangle - |u\uparrow u\downarrow d\uparrow\rangle - |u\downarrow u\uparrow d\uparrow\rangle\right]. \tag{4.102}$$

We've called the first line of Eq.(4.96) $|\Psi_{uud}\rangle$ (because each term has quark 1 u, quark 2 u, and quark 3 d) and then let the cyclic permutation operators act on this state. Remember that all these terms are necessary to make the state $|p\uparrow\rangle$ totally symmetric.

EXERCISE 4.51. Prove Eq.(4.101) by writing out the terms explicitly.

Now we can have another look at the expectation value in Eq.(4.99)

$$\mu_B = \langle p\uparrow|\hat{\mu}_B|p\uparrow\rangle \tag{4.103}$$

$$= \langle\Psi_{uud}|\,[1 + P_{123} + P_{132}]\,\hat{\mu}_B\,[1 + P_{123} + P_{132}]\,|\Psi_{uud}\rangle \tag{4.104}$$

$$= \langle\Psi_{uud}|\hat{\mu}_B|\Psi_{uud}\rangle + \langle\Psi_{uud}|P_{123}\hat{\mu}_B P_{132}|\Psi_{uud}\rangle + \langle\Psi_{uud}|P_{132}\hat{\mu}_B P_{123}|\Psi_{uud}\rangle \tag{4.105}$$

where, in the last line, we have used the fact that the off-diagonal matrix elements don't contribute.

EXERCISE 4.52. Argue (in other words, 'prove') than the 'direct' terms like $\langle\Psi_{uud}|P_{123}\hat{\mu}_B P_{123}|\Psi_{uud}\rangle$ are zero using the fact that $P_{123}P_{123} = P_{132}$.

The next thing we'll need to figure out are the expectation values

$$\langle\Psi_{uud}|P_{123}\hat{\mu}_B P_{132}|\Psi_{uud}\rangle \qquad \text{and} \qquad \langle\Psi_{uud}|P_{132}\hat{\mu}_B P_{123}|\Psi_{uud}\rangle. \tag{4.106}$$

What the heck are these? They're actually very simple. First, let's note something about the operator $\hat{\mu}_B$. Since it is the sum of three identical terms, it's unchanged (or 'invariant') under permutation of any of the particle label indices. If we write any permutation by the generic symbol P_π (where π can take on any of the six elements of the group S_3 (),(12),(13),(23),(123),(132)) then

$$[P_\pi, \hat{\mu}_B] = 0. \tag{4.107}$$

This means that any of the permutations *commutes* with the magnetic moment operator since it's a completely symmetric function with respect to the particle

labels. Obviously this holds for any operator which is a completely symmetric function of the particle labels, like the Hamiltonian. So you see that Eq.(4.107) is the mathematical expression of the PIP.

EXERCISE 4.53. Using the relation from Ex.(4.52), $P_{123}P_{123} = P_{132}$, show that $P_{123}P_{132} = \mathbb{1}$.

EXERCISE 4.54. Using Eq.(4.107) and the identity shown in the previous problem, show that

$$\begin{aligned} &\langle\Psi_{uud}|\hat{\mu}_B|\Psi_{uud}\rangle \\ =&\langle\Psi_{uud}|P_{123}\hat{\mu}_B P_{123}|\Psi_{uud}\rangle \\ =&\langle\Psi_{uud}|P_{132}\hat{\mu}_B P_{132}|\Psi_{uud}\rangle. \end{aligned} \tag{4.108}$$

EXERCISE 4.55. Finally, show:

$$\mu_B = 3\langle\Psi_{uud}|\hat{\mu}_B|\Psi_{uud}\rangle. \tag{4.109}$$

What we have done is show, in explicit detail, how Griffiths arrived at his solution for μ_p in his EXAMPLE 5.3 on P.181.

EXERCISE 4.56. Calculate the ratio of the magnetic moments for the proton and neutron: $\frac{\mu_n}{\mu_p}$.

8. Mass of the baryons

8.1. Discussion. We turn now to the calculation of the masses of the *light quark* baryons, N and Δ.[16] In the theory of relativity, we learn that the energy of an object (like an electron, proton, a horse, whatever) is related to its momentum $\mathbf{p}$ and its mass, m by

$$E = \sqrt{|\mathbf{p}|^2 + m^2}, \tag{4.110}$$

as Einstein instructed us all back in 1905. (The units in this equation are such that the speed of light, $c = 1$.) For an object at rest, $\mathbf{p} = 0$, we obtain the famous equation

$$E = mc^2 \tag{4.111}$$

with the speed of light, $c = 2.998 \times 10^8$ m/s now shown explicitly.

In quantum mechanics, we calculate the energy of an object moving with momentum, $\mathbf{p}$, by evaluating the Hamiltonian operator, $\hat{H}$ on the wave function, $|\Psi(\mathbf{p})\rangle$ which describes it

$$\hat{H}|\Psi(\mathbf{p})\rangle = E|\Psi(\mathbf{p})\rangle \tag{4.112}$$

$$= \sqrt{|\mathbf{p}|^2 + m^2}|\Psi(\mathbf{p})\rangle. \tag{4.113}$$

If the particle is at rest, we simply have

$$\hat{H}|\Psi\rangle = m|\Psi\rangle, \tag{4.114}$$

where we have written $|\Psi(\mathbf{p} = 0)\rangle = |\Psi\rangle$. The mass is then just the expectation value of the Hamilonian operator

$$m = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \tag{4.115}$$

In this case, the operator $\hat{H}$ is sometimes called the mass operator for obvious reasons.

Exercise 4.57. Prove this equation from Eq.(4.114). Write the expression for the mass, m in the case that the state $|\Psi\rangle$ is normalized to one.

What does this mass operator look like? Well, if we ignore interactions between the particles[17] then the energy of the system of particles is just the sum of their individual energies, similar to Eq.(4.2):

$$E = E_1 + E_2 + E_3. \tag{4.116}$$

We'll also make the apparently ridiculous assumption that the quarks aren't moving, so their *kinetic energy* is zero. It actually turns out that we can get away with making this assumption – but *not because it's true*! In fact, it couldn't be further from the truth. Perhaps we'll talk more about this point after we do the mass calculation.

[16]The light quarks are u and d (sometimes s – but not for our purposes).

[17]Obviously, there must be interactions between the particles. Otherwise they wouldn't be bound together.

Proceeding, we're assuming the quarks are at rest and not interacting. This means both the kinetic energy (energy of motion) and the potential energy (energy of position) are zero. The only other source of energy is the mass! So

$$E = m_1 + m_2 + m_3. \tag{4.117}$$

This means that the Hamiltonian is, obviously, $\hat{H} = \sum_i m_i$.

EXERCISE 4.58. Calculate the masses of the N and Δ states in terms of the mass of the light quarks m_q, where $m_u = m_d = m_q$ in the approximation that we neglect kinetic and potential energies of the quarks.

EXERCISE 4.59. In the approximation we're using (no kinetic, no potential, and $m_q = m_u = m_d$) prove that the masses of all the states: $p, n, \Delta^{++}, \Delta^{+}, \Delta^{0}, \Delta^{-}$ are all the same.

With this admittedly ridiculous assumption, you see from this Problem that the masses of the N and Δ are degenerate. This won't do, since their masses are observed to be[18]:

$$m_N = 939 \text{ MeV} \tag{4.118}$$

$$m_\Delta = 1232 \text{ MeV} \tag{4.119}$$

$$\delta_o \equiv m_\Delta - m_N = 293 \text{ MeV}. \tag{4.120}$$

There is a significant *observed mass splitting*, δ_o between the N and Δ states. Presently, we are calculating the mass splitting to be $\delta_c = 0$. How – if at all – can we reproduce the correct $\delta_c = \delta_o$ with our simple model?

There are two routes we can take. First, we might ask: Can we get the calculated mass splitting δ_c correct if we take into account the fact that the quarks are moving 'inside' the baryon? It turns out that this does give a significant contribution to δ_c if we take into account *spatial correlations* between the quarks. This is a might too ambitious at this stage since it involves using wave functions which have a factor, in addition to the spin, isospin, and color factors, which depends on the *positions* of the three quarks, $\psi_{pos}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$. This is very complicated, so let's try the other route.

What about potential energy? Can we get $\delta_c = \delta_o$ if we have some potential energies between the quarks? The answer, of course, is yes!

What sort of potential energy, which we'll call, $\hat{V}$ can we consider? We know that any charged particle with spin–$\frac{1}{2}$ has a magnetic moment. A magnetic moment generates a magnetic field in the vicinity of the particle. Another particle's magnetic moment, if close enough to be immersed in the magnetic field of the first particle, feels a force which tends to align it's magnetic moment (or spin) with the magnetic field. Therefore, if the spins of two quarks are aligned the force is attractive. If they're anti-aligned, it's repulsive.

Some of this information can be encoded into an operator which depends on the spins of the two quarks in question. The operator which gives the potential

[18]A word of caution by way of an excuse for my laziness: one should *always* include the precision to which an observed quantity is known. For example, the mass of the proton in MeV is $m_p = 938.27203(8) = 938.27203 \pm 0.00008$. This is one of the best known physical quantities and quite small, so we'll ignore the precision for it, along with the neutron mass which is known to the same precision. The Δ mass is much less well known (Why?) at (in MeV) $m_\Delta = 1232(2)$.

energy between quarks '1' and '2' is

$$V_{12} = V_{12}^0 \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2. \tag{4.121}$$

This is the so-called "hyperfine interaction." Here, the constant $V_{12}^0 \in \mathbb{R}$ is assumed to be $V_{12}^0 > 0$, and it will eventually be chosen to fit the observed data, *ie.* to make $\delta_c = \delta_o$. We should be careful here to mention that the form of the potential used here is a gross simplification of more sophisticated treatments of the force between pairs of quarks. In fact, there is no *potential* description of this force, in reality.[19] But experience has shown that this simple *model* (and extension) is useful in determining some of the properties of the nucleon. And it is very simple – it depends only on the spins of the quark pair, not their positions in space. We have 'convoluted' the spatial properties of the interaction into the constant V_{12}^0. Incidentally, the hyperfine interaction is *not* like the force between two bar magnets. For this force has a strong dependence on the *angular* orientation of this spins with respect to the line connecting them.

In any case, with the above caveats, we'll happily use this operator to calculate the quantity, δ_c.

Exercise 4.60. Using the results in Eqs.(4.27), show that if the two quark spins have $S = 1$ (ie. are aligned), the expectation value of the pair potential operator, $\hat{V}_{12}$ is positive. What is its value? Also calculate their potential energy when they have $S = 0$ (are anti-aligned) and show it is negative.

The total potential energy, $\hat{V}$ is simply the sum of the mutual interactions of all the pairs. In the case of three quarks then we have

$$\hat{V} = V_{12}^0 \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 + V_{13}^0 \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_3 + V_{23}^0 \boldsymbol{\sigma}_2 \cdot \boldsymbol{\sigma}_3 \tag{4.122}$$

$$= \sum_{i<j=1}^{3} V_{ij}^0 \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j. \tag{4.123}$$

The constants, V_{ij}^0 are allowed to depend on the properties of quarks i and j. For example, Griffiths chooses (in his Eq.(5.119))

$$V_{ij}^0 = \frac{A'}{4} \frac{1}{m_i m_j}. \tag{4.124}$$

Now we have our (more) complete Hamiltonian with which to calculate the mass splittings of the N and Δ:

$$\hat{H} = \sum_{i=1}^{3} m_i + \sum_{i<j=1}^{3} V_{ij}^0 \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j. \tag{4.125}$$

8.2. Mass calculation. We're in the home stretch now. We wish to calculate the masses of the N and the Δ which we refer to generically, as the state $|B, S_z\rangle$, where S_z is the total spin of the baryon projected on the z axis. The expectation

[19] The nucleon is a system of infinitely many bodies moving relativistically.

value we need is

$$m_B = \langle B, S_z|\hat{H}|B, S_z\rangle \tag{4.126}$$
$$= \sum_i m_i + \langle B, S_z|\hat{V}|B, S_z\rangle \tag{4.127}$$
$$= 3m_q + \sum_{i<j}\langle B, S_z|V^0_{ij}\boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j|B, S_z\rangle \tag{4.128}$$

Here, again, in the case of the proton (the Δ is much more simple and we leave the calculation of its mass as an exercise) we need to evaluate the $243 = 3 \times 9^2$ terms in the expectation value. So our first move is to check to see which matrix elements are zero by orthogonality. This time we have a little more complicated problem than in the evaluation of the magnetic moment, as the following problem demonstrates.

EXERCISE 4.61. Evaluate the off-diagonal matrix element

$$\langle u\uparrow u\uparrow d\downarrow|\boldsymbol{\sigma}_2 \cdot \boldsymbol{\sigma}_3|u\uparrow u\downarrow d\uparrow\rangle. \tag{4.129}$$

Apparently, we've got a big job here. Actually, we can simplify the problem significantly again. How? Exactly the same we exploited the *permutation symmetries* of the operator and wave functions in the case of the magnetic moment. Consider the term

$$\mathcal{M}_{12} = \langle B|\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2|B\rangle \tag{4.130}$$

where we've suppressed the S_z label in the wave function. By inserting appropriate permutation operators:

$$\mathcal{M}_{12} = \langle B|(\mathbb{1})\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2(\mathbb{1})|B\rangle \tag{4.131}$$
$$= \langle B|P_{ij}P_{ij}\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2P_{ij}P_{ij}|B\rangle. \tag{4.132}$$

EXERCISE 4.62. Complete the proof that $\mathcal{M}_{12} = \mathcal{M}_{13} = \mathcal{M}_{23}$ by choosing appropriate values of i and j for each equality and working out the algebra.

We've made a moderate simplification of the problem thus far. We now need to consider

$$m_B = 3m_q + (V^0_{12} + V^0_{13} + V^0_{23})\langle B, S_z|\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2|B, S_z\rangle. \tag{4.133}$$

This equation is general – it holds for any completely symmetric wave function $|B\rangle$.

EXERCISE 4.63. Prove Eq.(4.133).

This means we still have 81 terms for the proton. Uh-oh. Have no fear. We can further simplify. Let's write out the the proton wave function in it's explicit form which we get from Eq.(4.80), that is *before* we unpackage all the $|u\uparrow u\uparrow d\downarrow\rangle$ terms. We have

$$|p\uparrow\rangle = \frac{1}{\sqrt{2}}\Big\{|T;\tfrac{1}{2},+\tfrac{1}{2};\rho\rangle|S;\tfrac{1}{2},+\tfrac{1}{2};\rho\rangle + |T;\tfrac{1}{2},+\tfrac{1}{2};\lambda\rangle|S;\tfrac{1}{2},+\tfrac{1}{2};\lambda\rangle\Big\} \tag{4.134}$$
$$= \frac{1}{\sqrt{2}}\left\{|T;\tfrac{1}{2},+\tfrac{1}{2};\rho\rangle|S_{12};0,0\rangle|\uparrow\rangle\right.$$
$$\left.+|T;\tfrac{1}{2},+\tfrac{1}{2};\lambda\rangle\left[\sqrt{\frac{2}{3}}|S_{12};1,+1\rangle|\downarrow\rangle - \frac{1}{\sqrt{3}}|S_{12};1,0\rangle|\uparrow\rangle\right]\right\}. \tag{4.135}$$

This is the form we want to work with. What happens when we apply the operator $\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2$ to it? Let's work it out:

EXERCISE 4.64. Using the Eqs.(4.27) and realizing that each term in the above expression for $|p\uparrow\rangle$ has definite spin for particles '1' and '2', S_{12} evaluate the $\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 |p\uparrow\rangle$. Your answer should look a lot like Eq.(4.134) with different numerical coefficients in front of the terms with differing S_{12}.

EXERCISE 4.65. Using the results of the previous equation show that $\langle p\uparrow|\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2|p\uparrow\rangle = -1$.

Substitution into Eq.(4.133) gives

$$m_p = 3m_q - (V_{12}^0 + V_{13}^0 + V_{23}^0) \tag{4.136}$$

$$= 3m_q - \frac{3}{4}A' \frac{1}{m_q^2}, \tag{4.137}$$

where in the last line we have used Griffiths' form, Eq.(4.124) and assumed that $m_u = m_d = m_q$.

EXERCISE 4.66. Work out this equation and compare to Griffiths result in his Eq.(5.122).

EXERCISE 4.67. Calculate the mass of the Δ, m_Δ and compare with Griffiths formula Eq.(5.123).

EXERCISE 4.68. Calculate the mass splitting of the N and Δ, δ_c. Express the result as a function of m_q and A'.

EXERCISE 4.69. Using the numerical values for the quark masses and constant A' that Griffiths uses, calculate the values of the N and Δ masses. Use these to detemine δ_c and compare it to δ_o.