Group Theory in Quantum Mechanics: Part 2

The Lie Group $SU(2)$ and its connection to rotational transformations and angular momentum

Stefan Andreas Meyer

March 5, 2013

In Part 1 we learned about groups, their irreducible representations and how any representation can be block-diagonalized on invariant sub-spaces called isotypical components. In this part I'd like to introduce the Lie group SU(2) which is of central importance in Quantum Mechanics as it allows one to to describe rotational transformations of a system. With that it is possible to introduce the angular momentum operators as the generators of its Lie-Algebra which provides invaluable understanding of how to determine the total angular momentum of a system and for the meaning of the Clebsch Gordan coefficients.

At the end of this part I hope that the reader will have a better understanding of the following questions:

- What is a Lie Group, what's so special about it, and why is it that $SU(2)$ is the group of interest for Quantum Mechanics and not, say, $SO(3)$?
- What is a Lie Algebra, why are we interested in them and why is the angular momentum operator defined as the generator of the Lie-algebra $su(2)$?
- What are the irreducible representations for $SU(2)$ and how does that help us to determine the angular momentum of a Quantum Mechanical system?

1 Liegroups, Liealgebras & their relevance in QM

In order to make the formalism more tangible and easier to understand I will introduce the concept of a Liegroup by discussing the group $SO(3)$ which has particular relevance for a lot of physical problems as it is the set of rotations in $\mathbb{R}^3$. The concepts of $SO(3)$'s Liealgebra $so(3)$ will evolve rather naturally and show all the structure that even the beginner to Quantummechanics will already have seen in the context of the angular momentum operator. That this is no coincidence will be apparent by looking at a very simple example of a representation that describes rotations of the coordinate system for functions $\Psi \in L^2(\mathbb{R}^3)$.

1.1 Liegroups and Liealgebras

Let's begin by reminding ourselves of the definition of O(3) (Felder p.8):

$$O(3) \quad := \quad \left\{A \in \mathrm{Mat}(3, \mathbb{R}) : A^T A = \mathbb{1}\right\} \tag{1}$$

exercise: Show that $(O(3), \circ)$ is well-defined as a group[1];

With this definition it is easy to see that any $A \in O(3)$ introduces an **isometrie** (a map $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ for which $d(x,y) = d(f(x), f(y)$ holds true for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$): since $\langle A\mathbf{x}, \mathbf{y}\rangle = \sum_i(\sum_j A_{ij}x_j) \cdot y_i = \sum_j x_j \cdot (\sum_i A^T_{ji} y_i) = \langle \mathbf{x}, A^T\mathbf{y}\rangle$ it follows that $\langle A\mathbf{x}, A\mathbf{y}\rangle = \langle A^T(A\mathbf{x}), \mathbf{y}\rangle \overset{A\in O(3)}{=} \langle \mathbf{x}, \mathbf{y}\rangle$ which implies that $d(x,y)^2 := \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle = \langle A(\mathbf{x}-\mathbf{y}), A(\mathbf{x}-\mathbf{y})\rangle = \langle A(\mathbf{x}) - A(\mathbf{y}), A(\mathbf{x}) - A(\mathbf{y})\rangle = d(A(\mathbf{x}), A(\mathbf{y}))^2$. It is interesting to see that apart from translations *all* isometries are induced by orthogonal matrices (Felder p.43)

We can now introduce $SO(3)$ as a subgroup of $O(3)$:

$$SO(3) \quad := \quad \{A \in O(3) : \det(A) = 1\} \tag{2}$$

exercise: Show that $SO(3)$ is well-defined[2]

In the exercise to show that $O(3)$ is well defined, we already saw that $\det(A) = \pm 1$ for all $A \in O(3)$. Consequently any $A \in O(3)$ is either already in the subgroup $SO(3)$ or (if $\det(A) = -1$) it can be written as a product $A = (-\mathbb{1})(-\mathbb{1})A = (-\mathbb{1})(-A)$ of the inversion matrix $-\mathbb{1}$ and a matrix $-A \in SO(3)$. Now we would like to show that all elements of $SO(3)$ can be interpreted as rotations.

[1] we need to show a) that $O(3)$ contains $\mathbb{1}$ (obvious); b) that $A \in O(3)$ mandates $A^{-1} \in O(3)$ (this can be seen since $1 = \det(\mathbb{1}) = \det(A^T A) = \det(A^T) \cdot \det(A) = \det(A)^2 \implies \det(A) \neq 0 \implies A$ is invertible and $A^{-1} = A^T \implies (A^{-1})^T A^{-1} = (A^T)^T A^{-1} = AA^{-1} = \mathbb{1}$) c) that $A, B \in O(3)$ mandates $AB \in O(3)$ (this is straight forward since $(AB)^T AB = B^T A^T AB = B^{-1}A^{-1}AB = \mathbb{1}$)

[2] again we need to show a) that $SO(3)$ contains $\mathbb{1}$ (obvious); b) that $A \in SO(3)$ mandates $A^{-1} \in SO(3)$ (since $\det(A) \neq 0 \implies A$ is invertible and $A^{-1} = A^T \implies \det(A^{-1}) = \det(A^T) = \det(A) = 1$) c) that $A, B \in SO(3)$ mandates $AB \in SO(3)$ (this is straight forward since $\det(AB) = \det(A) \cdot \det(B) = 1$)

It is a simple Corrolary of the the Normal-Form Theorem in linear Algebra (Gerd Fischer, "Lineare Algebra", page 292) that any matrix $A \in SO(3)$ can be written as:

$$A = \overbrace{O \underbrace{\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=:R(\mathbf{e}_3,\theta)} O^{-1}}^{=:R(\mathbf{n},\theta)} \quad \text{with} \quad O \in SO(3) \tag{3}$$

It is obvious that the matrix $R(\mathbf{e}_3, \theta)$ corresponds to a rotation around $\mathbf{e}_3$ by the angle θ and since A is equivalent to $R(\mathrm{e}_3, \theta)$ with a new orthonormal basis $\mathcal{B} = (O\mathbf{e}_1, O\mathbf{e}_2, O\mathbf{e}_3)$ it corresponds to a rotation around $\mathbf{n} := O\mathbf{e}_3$.

Up to this point we have learned:

- that all $A \in O(3)$ induce canonically isometries on $\mathbb{R}^3$;
- that all $A \in O(3)$ are already in $SO(3)$ themselves or can be written as a product of an inversion and an element of $SO(3)$;
- that any $A \in SO(3)$ can be interpreted as a rotation $R(\mathbf{n}, \theta)$ by an angle θ around the axis $\mathbf{n}$.

This is already a lot but something quite obvious is missing: when we look at the function $R(\mathbf{e}_3, \cdot) : \mathbb{R} \longrightarrow SO(3)$ it is apparent that each matrix-element $R(\mathbf{e}_3, \cdot)_{ij} : \mathbb{R} \longrightarrow \mathbb{R}$ is continously differentiable and even an element of $C^\infty(\mathbb{R})$. But in order to interpret $R(\mathbf{e}_3, \cdot)$ as a differentiable function we need a norm (or at least a metric) on $SO(3)$. This norm can be defined canonically by identifying $\mathrm{Mat}(n, \mathbb{K})$ with $\mathbb{K}^{n^2}$ (here $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$):

$$\Phi : \mathrm{Mat}(n, \mathbb{K}) \longrightarrow \mathbb{K}^{n^2} \tag{4}$$

$$X \longmapsto (X_{11}, X_{12}, ..., X_{nn}) \tag{5}$$

Since for any $x \in \mathbb{K}$ the norm is given by $\|x\| = \sqrt{\sum |x_i|^2}$ it makes sense to define $\|X\| = \sqrt{\mathrm{tr}(X^*X)}$ as **the norm of a matrix** $X \in \mathrm{Mat}(n, \mathbb{K})$.

exercise: a) show that this definition of the norm coincides with the norm on $\mathbb{K}^{n^2}$ in effect show that $\|X\| = \|\Phi(X)\|$[3]; b) show the validity of $\|XY\| \leq \|X\|\|Y\|$ (called the Schwarz inequality)[4]

With this norm and everything else we have learned so far it is obvious ...

---

[3] $(X^*X)_{ij} \stackrel{\text{Def.}}{=} (\overline{X^T}X)_{ij} = \sum_k \overline{X^T}_{ik} X_{kj} = \sum_k \overline{X}_{ki} X_{kj} \implies \mathrm{tr}(X^*X) \stackrel{\text{Def.}}{=} \sum_i (X^*X)_{ii} = \sum_{i,k} \overline{X}_{ki} X_{ki} = \sum_{i,k} |X_{ki}|^2 = \|\Phi(X)\|^2$

[4] the norm squared of the matrix XY is the sum of the norm squared of its elements: $\|XY\|^2 \stackrel{\text{see a)}}{=} \sum_{i,j} |(XY)_{ij}|^2 = \sum_{i,j} |\sum_k X_{ik} Y_{kj}|^2$; if we now define $u, v \in \mathbb{K}^n$ for fixed i and j as $u_k := \bar{X}_{ik}$ and $v_k := Y_{kj}$ then we can use the Schwarz inequality $|\langle u, v\rangle|^2 \leq \|u\|^2 \|v\|^2$ to obtain the assertion;

(a) ... that $R(\mathbf{e}_3, \cdot)$ is continously differentiable on $(\mathrm{Mat}(n, \mathbb{K}), \|\cdot\|)$;

(b) ... that $R(\mathbf{e}_3, 0) = \mathbb{1}$;

(c) ... that $R(\mathbf{e}_3, \theta + \phi) = R(\mathbf{e}_3, \theta) R(\mathbf{e}_3, \phi)$;

A map with such characteristics is more generally called a **one parameter group**.

Now it is easy to make a Taylor expansion of $R(\mathbf{e}_3, \cdot) \in C^\infty(\mathrm{Mat}(n, \mathbb{K}))$ which yields

$$R(\mathbf{e}_3, \theta) = \underbrace{\mathbb{1}}_{R(\mathbf{e}_3,0)} + \theta \cdot \underbrace{\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\partial_\theta R(\mathbf{e}_3,0)} + \frac{1}{2}\theta^2 \cdot \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\partial_\theta^2 R(\mathbf{e}_3,0)} + ... \tag{6}$$

<u>*exercise:*</u> a) Show that for $n \geq 1$ the dervatives are given by: $\partial_\theta^{n+2} R(\mathbf{e}_3, 0) = -\partial_\theta^n R(\mathbf{e}_3, 0)$.
b) Show that for $I_3 := \partial_\theta R(\mathbf{e}_3, 0)$ the product is given by $I_3 I_3 = \left(\begin{array}{c|c} -\mathbb{1} & 0 \\ \hline 0 & 0 \end{array}\right)$

From the previous exercise it is apparent that the Taylor expansion of $R(\mathbf{e}_3, \cdot)$ in qu. 6 coincides with the definition of the exponential function

$$R(\mathbf{e}_3, \theta) = \exp(\theta I_3) \tag{7}$$

where I_3 is called the **infinitesimal generator** of the one-parameter group. On a practical level one could wonder now if $\exp(X)$ exists for all $X \in \mathrm{Mat}(n, \mathbb{K})$ (i.e. if it always converges) and if it handles well (i.e. if it respects the "normal" operation-rules which we are accustomed to from $X \in \mathbb{R}$. On a more fundamental level we could wonder if $\exp(\theta X)$ is *always* a one-parameter group and if all parameter groups are of that form. As it turns out the answer to all those questions is positive and can be found in Felder p.59 - p.61[5].

Simply by knowing I_3 and rotating the axises in a cyclic fashion, it is easy to see that the infinitesimal generators for the rotations $R(\mathbf{e}_1, \theta), R(\mathbf{e}_2, \theta)$ and $R(\mathbf{e}_3, \theta)$ are as follows:

$$I_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad I_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad I_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{8}$$

One could wonder now if any rotation $R(\mathbf{n}, \theta)$ can be written as $\exp(\theta X)$ and, if so, how exactly the infinitesimal generator X looks like. The answer to the first question is positive

[5] For $X, Y \in \mathrm{Mat}(n, \mathbb{K})$ it is possible to show the following: a) $\exp(-X) = (\exp X)^{-1}$ (particularly X is invertible); b) $\exp(X^*) = (\exp(X))^*$ (particularly $\exp(X^T) = (\exp(X))^T$); c) $A\exp(X)A^{-1} = \exp(AXA^{-1})$ d) $\det(\exp(X)) = \exp(\mathrm{tr}(X))$ and e) $\exp(X)\exp(Y) = \exp(X + Y)$ if $XY = YX$

and the expression for $X_{\mathbf{n}}$ turns out to be surprisingly simple:

$$R(\mathbf{n},\theta) = \exp(\theta \underbrace{\sum n_i I_i}_{X_{\mathbf{n}}}) \tag{9}$$

<u>*exercise:*</u> Show that the above relation is true by showing that a) $\exp(\theta X_{\mathbf{n}})^T \exp(\theta X_{\mathbf{n}}) = \mathbb{1}$ which implies trivially that $\exp(\theta X_{\mathbf{n}})$ is an orthogonal matrix (i.e. an element of $O(3)$)[6]; b) $\det(\exp(\theta X_{\mathbf{n}})) = 1$ which implies again trivially that $\exp(\theta X_{\mathbf{n}})$ is in $SO(3)$ and therefore a rotation[7]; c) $X_{\mathbf{n}}\mathbf{n} = 0$ which implies that $\exp(\theta X_{\mathbf{n}})\mathbf{n} = \mathbf{n}$ and therefore that $\mathbf{n}$ is the axis of rotation[8]; d) that θ is indeed the angle of roation[9]

This is an important result not only on a practical level but also more fundamentally. It means that the set $so(3) := \{X \in \mathrm{Mat}(n,\mathbb{K}) : \exp(\theta X) \in SO(3) \quad \text{for all } \theta \in \mathbb{R}\}$ has the structure of a real vectorspace V spanned by I_1, I_2 and I_3. Furthermore it is easy to see that the operation

$$[\cdot,\cdot]: \quad so(3) \times so(3) \longrightarrow so(3) \tag{10}$$

$$(A,B) \longmapsto [A,B] := AB - BA \tag{11}$$

is well defined and adds more structure to V with the following properties

(i) $[\lambda X + \mu Y, Z] = \lambda[X,Z] + \mu[Y,Z]$ for $\lambda,\mu \in \mathbb{K}$

(ii) $[X,Y] = -[Y,X]$

(iii) $[[X,Y],Z] + [[Z,X],Y] + [[Y,Z],X] = 0$ (called Jacobi identity)

<u>*exercise:*</u> Show that $[\cdot,\cdot]$ is well defined by verifying that $[I_i, I_j] = \epsilon_{ijk} I_k$ and that furthermore properties (i)-(iii) are correct.

It turns out that the structure which we just discovered for $so(3)$ is more general: Any real or complex vector space g equipped with a "**Lie-bracket**" $[\cdot,\cdot] : g \times g \longrightarrow g$ that has the properties (i)-(iii) is called a **Lie-Algebra**. A Homomorphism of a Lie-algebra therefore not only respects addition (i.e. is not only linear) but also respects the Lie-bracket.

Furthermore it is possible to show that for any closed[10] subgroup of $GL(n,\mathbb{K})$ (referred to more generally as **(Matrix-) Lie-group**) the set

$$g \equiv \mathrm{Lie}(G) = \{X \in \mathrm{Mat}(n,\mathbb{K}) : \ \exp(\theta X) \in G \quad \text{for all } \theta \in \mathbb{R}\} \tag{12}$$

[6] $\exp(\theta X_{\mathbf{n}})^T \exp(\theta X_{\mathbf{n}}) = \exp(\theta X_{\mathbf{n}}^T + \theta X_{\mathbf{n}}) = \exp(\theta(-X_{\mathbf{n}}) + \theta X_{\mathbf{n}}) = \exp(0) = \mathbb{1}$ since $I_j^T = -I_j$

[7] $\det(\exp(\theta X_{\mathbf{n}})) = \exp(\mathrm{tr}(\theta X_{\mathbf{n}})) = 1$ since $\mathrm{tr}(I_j) = 0$ for all j;

[8] to show that $X_{\mathbf{n}}\mathbf{n} = 0$ one simply has to use the definition for $X_{\mathbf{n}} = \sum n_i I_i$ and the definition for I_i; the rest is evident

[9] let's look at $\mathbf{n} = (1,0,0)$; then by simple calculation and using the powerseries expression for exp, sin and cos one sees that $R(\mathbf{n};\theta)\mathbf{e}_3 = -\sin(\theta)\mathbf{e}_2 + \cos(\theta)\mathbf{e}_3$;

[10] closed with respect to the topology introduced by the metric;

has the structure of a Lie-algebra and is called the **Lie-algebra of the Lie-group** G.

<u>*exercise:*</u> Show that for the following examples are correct[11] (Felder page 62):

$$\begin{aligned} \mathrm{gl}(n,\mathbb{K}) &\overset{a}{=} \mathrm{Mat}(n,\mathbb{K}) && (13)\\ u(n) &\overset{b}{=} \{X \in \mathrm{Mat}(n,\mathbb{C}):\ X^* = -X\} && (14)\\ su(n) &\overset{c}{=} \{X \in \mathrm{Mat}(n,\mathbb{C}):\ X^* = -X\ \ \mathrm{tr}X = 0\} && (15)\\ o(n) = so(n) &\overset{d}{=} \{X \in \mathrm{Mat}(n,\mathbb{C}):\ X^T = -X\} && (16) \end{aligned}$$

This already shows a very important aspect of Lie-Algebras: while it can be true that $\exp(g) = G$ it doesn't have to be (as in the case of $O(3)$ where I can only reach the Zusammenhangskomponente of $\mathbb{1}$). In general I can only reconstruct the vicinity of $\mathbb{1}$.

1.2 Representations of Liegroups

In the previous sections we introduced the definition of a Liegroup and its Liealgebra by discussing $SO(3)$ and its connection to $so(3)$. It is important to see that the special feature about a Liegroup is the fact that its group-structure is compatible with its topology (induced by the norm), a fact that becomes apparent when looking at the one parameter groups that let us describe rotations about an axis. In this context it is only natural that we demand of any **representation** ρ **of a Liegroup** $(G, \circ)$ on a vectorspace V

$$\rho: \quad G \longrightarrow GL(V) \tag{17}$$

to not only respect the group-structure (i.e. to be a Homomorphism) but also to respect the group's topology (i.e. to be continous). This will allow us now to show that the image $\psi_t := \rho(\phi_t) \in \mathrm{GL}(V)$ of a one-parameter group ϕ_t with ρ is again a one-parameter group. This is straight forward to see:

(i) $\psi_{t+s} = \psi_t\psi_s$ follows directly from ϕ being a one parameter group and ρ being a Lie group homomorphism.

(ii) $\psi_0 = 1$ follows similarly.

(iii) ψ_t is continous because ρ is. If we now assume for a moment that $\dot{\psi}_0$ exists[12] then

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_t = \lim_{h\to 0}\frac{\psi_{t+h} - \psi_t}{h} = \psi_t \lim_{h\to 0}\frac{\psi_h - \psi_0}{h} = \psi_t\,\dot{\psi}_0 \tag{18}$$

[11] **ad a)** for any $X \in \mathrm{Mat}(n,\mathbb{K})$ we know that $\det(\exp(\theta X)) = \exp(\theta\ \mathrm{tr}(X)) \neq 0 \Longrightarrow \exp(\theta X)$ is invertible and therefore in $\mathrm{GL}(n,\mathbb{K})$; **ad b)** for any $X \in u(n)$ we know by definition that $(\exp(\theta X))^* \exp(\theta X) = \mathbb{1}$; using $(\exp(\theta X))^* = (\exp(\theta X^*))$ and the product-rule, derivation $\frac{\mathrm{d}}{\mathrm{d}\theta}\big|_{\theta=0}$ yields $X^* + X = 0$; **ad c)** for any $X \in su(n)$ we know furthermore (also by definition) that $\det(\exp(\theta X)) = 1 \implies \exp(\theta\ \mathrm{tr}(X)) = 1 \implies \mathrm{tr}(X) = 0$ **ad d)** note that for $X \in \mathrm{Mat}(n,\mathbb{R})$ the condition $X^* = -X$ automatically implies that all diagonal elements of X vanish $(X_{ii} = 0) which means that$ tr(X)=0 and therefore $so(3) = o(3)$

[12] for a detailed proof see Felder p.69

and therefore ψ_t is even continouly differentiable.

The pervious considerations allow us now to seek a definition of the **representation** ρ^* **of the Liealgebra** $(\mathrm{Lie}(G), \circ)$ that is sensible (i.e. that makes the following diagram commutative):

$$\begin{array}{ccc} G & \xrightarrow{\rho} & \mathrm{GL}(V) \\ {\scriptstyle \exp}\uparrow & & \uparrow{\scriptstyle \exp} \\ \mathrm{Lie}(G) & \xrightarrow[\rho^*]{} & \mathrm{gl}(V) \end{array} \tag{19}$$

To this end let us consider $\psi_t = \rho(\exp(tX))$ with $t \in \mathbb{R}$ and $X \in \mathrm{Lie}(G)$. From the pervious paragraph we know that ψ_t is a one-parameter group in $\mathrm{GL}(V)$ and from the previous section that every one-parameter group has the form $\exp(tY)$ with $Y \in \mathrm{Lie}(\mathrm{GL}(V)) \equiv \mathrm{gl}(V)$:

$$\rho(\exp(tX)) = \exp(tY) \tag{20}$$

$$\Longrightarrow \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho(\exp(tX)) = \underbrace{\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp(tY)}_{=Y} \tag{21}$$

Inserting eq.(21) into eq.(20) yields:

$$\rho(\exp(tX)) = \exp\left(t\left[\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho(\exp(tX))\right]\right) \tag{22}$$

This is the motivation to define

$$\rho^*(X) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho(\exp(tX)) \tag{23}$$

since it renders diagram (19) commutative! Fortunately this definition also respects the Lie algebra's structure such that

a) ρ^* is linear: $\rho^*(\lambda X + \mu Y) = \lambda\, \rho^*(X) + \mu\, \rho^*(Y)$[13];

b) $\rho*$ preserves the Lie-bracket: $\rho^*([X, Y]) = [\rho^*(X), \rho^*(Y)]$[14];

In general we shall demand that any representation of a Lie algebra have these properties.

This brings us to the **main theorem** of this section: that a representation ρ on a vector space V of a simply connected group is irreducibel if and only if ρ^* is irreducibel[15].

[13] Felder, page 70

[14] Felder, page 70

[15] *Felder page 69 and 70*

Proof: Let $W \subseteq V$ be a closed subvectorspace not equal to $\{0\}$ and $\gamma_t : \mathbb{R} \longrightarrow W$ and γ_t a one-parameter group on W. This means that $\frac{\mathrm{d}\gamma_t}{\mathrm{d}t} \in W$, too, and especially $\frac{\mathrm{d}\gamma_t}{\mathrm{d}t}(t=0)$. Let now ρ be invariant on W and $\gamma_t := \rho(\exp(tX))(w)$ for $w \in W$ arbitrary but fixed. Then it is obvious that

$$\rho^*(X) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \underbrace{\rho(\exp(tX))}_{\gamma_t} \tag{24}$$

has W as an invariant subspace, too. On the other hand, if ρ^* is invariant on W (i.e. $\rho^*(X)(w) \in W$ for all $w \in W$ and for all X in the group's Liealgebra). Then

$$\rho(\exp(tX))(w) \stackrel{\text{Def.}}{=} \exp(t\rho^*(X))(w) = \sum_j \frac{t^j}{j!}\rho^*(X)^j(w) \ \in \ W \tag{25}$$

So a non-trivial invariant subspace W of ρ is also one of ρ^* and vice versa. In order to infer the irreducibility of ρ^* from ρ (and vice versa) such a subspace W also needs to be minimal (i.e. it may not contain a smaller, non-trivial invariant subspace $\tilde{W} \subsetneq W$). But that is easy to see: let us assume that ρ is irreducible on W but ρ^* is not (i.e. W is minimal for ρ but not minimal for ρ^*). Then there is a non-trivial subvector space $\tilde{W} \subsetneq W$ on which ρ^* is irreducible. Because of our previous considerations, $\tilde{W}$ would also be invariant under ρ which would be a contradiction to our initial assumption that ρ is irreducible on W. With that the theorem is proven.

1.3 Practical example of a Liegroup representation in QM

The use of the above theorem becomes quickly apparant when looking at the simple example of describing the rotation of coordinates $x \mapsto \tilde{x} = Rx, \quad R \in SO(3)$ for a wavefunction $\Psi \in L^2(\mathbb{R}^3)$. This can be expressed most conveniently with a representation ρ:

$$\rho : \quad SO(3) \longrightarrow \mathrm{GL}(L^2(\mathbb{R}^3)) \tag{26}$$

$$R \longmapsto \rho_R \ \text{ with } \ (\rho_R\Psi)(x) := \Psi(R^{-1}x) \tag{27}$$

<u>*exercise:*</u> show that this definition is indeed a homomorphism i.e. that $\rho_{R\tilde{R}} = \rho_R \circ \rho_{\tilde{R}}$;[16]

If we now look at $\rho^* : \quad so(3) \longrightarrow \mathrm{gl}(L^2(\mathbb{R}^3))$ we can make an interesting connection to the

[16] it just an exercise in writing it out; the 2 things important to remember/ see are a) that $(\rho_R(\rho_{\tilde{R}}(\Psi)))(x) = (\rho_R(\Psi(\tilde{R}^{-1}\cdot)))(x) = \Psi(\tilde{R}^{-1}(R^{-1}x))$ and b) $(\rho_{R\tilde{R}}\Psi)(x) = \Psi(\tilde{R}^{-1}(R^{-1}x))$ since $(R\tilde{R})^{-1} = \tilde{R}^{-1}R^{-1}$

angular momentum operator L in Quantum mechanics:

$$\rho^*_{X_\mathbf{n}}\Psi(\mathbf{x}) \overset{\text{Def}\,\rho^*}{=} \left.\frac{\mathrm{d}}{\mathrm{d}\alpha}\right|_{\alpha=0} \left\{\rho_{\exp(\alpha X_\mathrm{n})}\Psi(\mathrm{x})\right\} \tag{28}$$

$$\overset{\text{Def}\,\rho}{=} \left.\frac{\mathrm{d}}{\mathrm{d}\alpha}\right|_{\alpha=0} \{\Psi(\exp(-\alpha X_\mathrm{n})\mathrm{x})\} \tag{29}$$

$$\overset{\text{chain rule}}{=} (\nabla\Psi)\circ(\exp(-\alpha X_\mathrm{n})(-X_\mathbf{n})\mathbf{x})|_{\alpha=0} \tag{30}$$

$$\overset{\exp(0)=\mathbb{1}}{=} -(\nabla\Psi)\circ(X_\mathbf{n}\mathbf{x}) \tag{31}$$

$$\overset{X_\mathbf{n}\mathbf{x}=\mathbf{n}\wedge\mathbf{x}}{=} -(\nabla\Psi)\circ(\mathbf{n}\wedge\mathbf{x}) \tag{32}$$

$$\overset{\rho}{=} -\mathbf{n}\circ(\mathbf{x}\wedge\nabla\Psi) \tag{33}$$

$$\overset{\rho}{=} -\frac{i}{\hbar}\mathbf{n}\circ\underbrace{\frac{\hbar}{i}(\mathbf{x}\wedge\nabla)}_{(L_x,L_y,L_z)}\Psi \tag{34}$$

This means in particular that

$$L_i = i\hbar\,\rho^*_{I_i} \tag{35}$$

which means in turn that the Lie-bracket is passed on (*Straumann page 144 and 145*):

$$[L_i, L_j] = i\hbar\sum_k \epsilon_{ijk}L_k \tag{36}$$

1.4 Finding the angular momentum using group theory

In the last two subsections we had two important results:

(a) The angular momentum operators L_i are induced through the representation ρ by the relation $L_i = i\hbar\,\rho^*_{I_i}$.

(b) For a continous group like $SO(3)$ the representation ρ is irreducible if and only if ρ^* is irreducible.

This will allow us now to show that the Isotypical components $W_{D^k} = V^1_{D^k}\oplus V^2_{D^k}\oplus ... \subseteq L^2(\mathbb{R}^3)$ of ρ are already the Eigen-spaces $\mathrm{Eig}(\mathbf{L}^2,\lambda)$ of $\mathbf{L}^2 := \sum_k L_k$ (here λ denotes the eigen-value). Since we have already developed a lot of tools to find such invariant subspaces of a representation, this will be very helpful.

Let's start by having a look at $so(3) = \{X\in \mathrm{Mat}(n,\mathbb{C}):\ X^T = -X\}$ and its generators I_k. It takes a bit of calculation[17] to show that:

$$[\sum L_k^2, \rho^*_g] = 0 \tag{37}$$

[17] Since any $g\in so(3)$ can be expressed as a linear combination of the infinitesimal generators I_k, and since furthermore ρ^* and the Liebracket are linear, it suffices to show that $[\sum L_k^2, \rho^*_{I_k}] = 0$; and

If we now consider the invariant subspace $V^i_{D^k} \subseteq L^2(\mathbb{R}^3)$ on which $\rho|_{V^i_{D^k}}$ is irreducible (and equivalent to the irreducible representation D^k of $SO(3)$) then we know that $\rho^*|_{V^i_{D^k}}$ is also irreducible and it follows straight from Schur's Lemma[18] that $\mathbf{L}^2 = \lambda \cdot \mathrm{id}$. Using a result that we will proof further down (section 4.1) we can even say that $\mathbf{L}^2 = k(k+1) \cdot \mathrm{id}$ and hence $W_{D^k} \subseteq \mathrm{Eig}(\mathbf{L}^2, k(k+1))$.

We can see that $\mathrm{Eig}(\mathbf{L}^2, k(k+1)) \subseteq W_{D^k}$ by contradiction: lets pick a $v \neq 0$ with $\mathbf{L}^2(v) = k(k+1)v$ which we assume to be in a different isotypical component $W_{D^{\tilde{k}}}$ with $k \neq \tilde{k}$. Since we know from above that $W_{D^{\tilde{k}}} \subseteq \mathrm{Eig}(\mathbf{L}^2, \tilde{k}(\tilde{k}+1))$ our assumption leads to a contradiction. We can therefore conclude that there is no $v \neq 0$ in $\mathrm{Eig}(\mathbf{L}^2, k(k+1))$ which is in an isotypical component $W_{D^{\tilde{k}}}$ with $k \neq \tilde{k}$.

since $\rho^*_{I_k} = -\frac{i}{\hbar} L_k$ it is actually sufficient to show that $[\sum L_k^2, L_j] = 0$ for $j = 1, 2, 3$; Let us do the calculation for j=1: $[\sum L_k^2, L_1] = [L_1^2, L1] + [L_2^2, L_1] + [L_3^2, L_1] = 0 + (L_2^2 L_1 - L_1 L_2^2) + \text{third term} = 0 + [(L_2(L_2 L_1) - L_2(L_1 L_2)) + ((L_2 L_1) L_2 - (L_1 L_2) L_2)] + \text{third term} = 0 + (L_2[L_2, L_1] + [L_2, L_1] L_2) + \text{third term} = 0 - (L_2 L_3 + L_3 L_2) + (L_3 L_2 + L_2 L_3) = 0$

[18]note that $\rho^*_g(V^i_{D^k}) \subseteq V^i_{D^k}$ for all $g \in so(3)$ and therefore especially for $g = \sum I_k^2$; hence $V^i_{D^k}$ is an invariant subspace for both, ρ^* and $\mathbf{L}^2$ and Schur's Lemma is applicable

2 Formalising symmetry considerations

Symmetry is a word that is commonly used in our every-day language and we all have some kind of an intuitive notion of what we think it means leaving the door wide open for misunderstandings when referring to it in a more formal, scientific context. This is especially true in the context of Quantum Mechnaics where physical objects are described (rather counter intuitively) by rays in a separable Hilbertspace and observables by self-adjoint operators. In order to come up with a sensible definition for the word "symmetry" we need to ask the question what we actually want from a symmetry and we need to find a way to express it mathematically.

In some cases it might seem straight forward: if, for example, the vectorspace for describing our quantum-mechanical system is $L^2(\mathbb{R}^3)$ and the potential is rotation invariant (e.g. $1/r$), then it might seem obvious that the representation $\rho: \quad SO(3) \longrightarrow \mathrm{GL}(L^2(\mathbb{R}^3))$ with $\rho_R \Psi(x) := \Psi(R^{-1}x)$ which describes rotations around the origin is also a symmetry transformations of this system (even though we haven't really specified what we mean by symmetry yet). But already in this simple example of a $SO(3)$ representation it is easy to show that it is not possible to describe a particle with half-integer spin like the electron[19]. For the latter the proper Hilbert-space is $L^2 \bigotimes \mathbb{C}^2$ and the group of choice for describing rotations is $SU(2)$. Here it might not be apparent right away how to express a rotation.

One could also ask if a given representation ρ really describes *all* the symmetries that the system has or if there are maybe more that are hidden or just not that obvious. This is a valid question because even in classical mechanics the $1/r$-potential-problem has a $SU(4)$ rather than just $SO(3)$ symmetry reflecting that not only the angular momentum but also of the Runge-Lenz-vector is conserved[20].

All this makes it apparent that it is necessary to formalize symmetry-considerations which shall be done in the following: first we quickly review some basic terms like state, observable and probability measure. Then we will discuss Automorphisms in general and finally define a symmetry as a special case of an Automorphism.

2.1 States, observables and probability measure

In Quantum mechanics the pure state of a system is given by a ray in a separable Hilbert Space H which is an equivalence class $[\Psi]$ of vectors in H that only differ by a phase factor: $\Psi \sim \Phi \iff \Psi = e^{i\phi}\Phi$. Its temporal evolution is given by

$$\Psi(t) = \exp\left(-\frac{i}{\hbar}Ht\right)\Psi_0 \tag{38}$$

[19] For more information one can read up on the Stern-Gerlach experiment.

[20] In quantum mechanics we also have a $SU(4)$ symmetry resulting in a higher degeneracy of the energy levels than would be the case if it was just a $SO(3)$ symmetry

where H is the Hamilton operator of the system. Obviously $\Psi(t)$ solves the Schroedinger equation and it is straight forward to see that $U(t) \equiv \exp\left(-\frac{i}{\hbar}Ht\right)$ is a unitary operator[21].

Once we know Ψ_0 and H we can answer the question what the probability is to measure a certain observable A in a set $\Delta \subseteq \mathbb{R}$ by calculating the probability measure

$$W^A_{[\Psi]}(\Delta) := \langle \Psi, E^A(\Delta)\Psi \rangle \tag{39}$$

where $E^A(\Delta)$ is the projection valued measure of the self-adjoint operator A[22]. In order to make this expression more tangible let us look at the examples of an harmonic oscillator where it is easy to write down $W^A_{[\Psi]}(\Delta)$ explicitly in a manner that is easily recognized.
The harmonic oscillator's Hamilton operator has a discrete spectrum of Eigen-values and can therefore be written as $H = \sum_i \lambda_i |v_i\rangle\langle v_i|$ in Dirac notation. The probability $W^H_{[\Psi]}(\Delta)$ to measure its energy in the interval $\Delta \subseteq \mathbb{R}$ is given by:

$$W^H_{[\Psi]}(\Delta) = \sum_{i \text{ with } \lambda_i \in \Delta} \langle \Psi \underbrace{|v_i\rangle\langle v_i|}_{E^H(\Delta)} \Psi \rangle \tag{40}$$

$$= \sum |\alpha_i|^2 \quad \text{where } \Psi = \sum \alpha_i |v_i\rangle \tag{41}$$

which is the expression we are used to from introductory Quantum mechanics.

2.2 Automorphisms

When we describe a quantum mechanical system by specifying its state $[\Psi]$ and its observables $A, B, C, ...$ it is only natural to ask what kind of maps

$$\alpha : [\Psi] \longrightarrow [\tilde{\Psi}] \tag{42}$$

$$A \longrightarrow \tilde{A} \tag{43}$$

will leave the probability measures (and that's the only thing relevant for comparing theory and experiment) unmodified:

$$W^A_{[\Psi]}(\Delta) = W^{\alpha(A)}_{\alpha([\Psi])}(\Delta) \tag{44}$$

An surjective α with this property is called an **Automorphism** and it is easy to see that any unitary operator U will induce such an Automorphism:

$$W^{UAU^{-1}}_{[U\Psi]}(\Delta) = \langle U\Psi, UE^A(\Delta)U^{-1}U\Psi \rangle \tag{45}$$

$$= \langle U^*U\Psi, E^A(\Delta)\Psi \rangle \tag{46}$$

$$\overset{U^*=U^{-1}}{=} \langle \Psi, E^A(\Delta)\Psi \rangle = W^A_{[\Psi]}(\Delta) \tag{47}$$

[21] $U(t)^* = \left[\exp\left(-\frac{i}{\hbar}Ht\right)\right]^* = \exp\left(-\frac{i}{\hbar}Ht\right)^* = \exp\left(\frac{i}{\hbar}H^*t\right) \overset{H=H^*}{=} U(t)^{-1}$

[22] According to the spectral theorem such a measure exists for every self-adjoint operator A and $E^A(\mathbb{R}) = \text{id}$, for more details see the book by Eduard Prugovecki "Quantum Mechanics in Hilbert-Spaces"

But it is also possible to show that *any* Automorphism can be induced by a unitary operator which is unique safe a phase factor. This means that if $\alpha([\Psi]) = [U\Psi] = [\tilde{U}\Psi]$ then it follows that $\tilde{U} = e^{i\theta}U$ with $\theta \in \mathbb{R}$.

2.3 Symmetries

With the general remarks of the previous two subsections in mind we can now not only define what a symmetry is, but also see the sense behind its definition: let ρ be the representation of a group G onto the set of Automorphisms (i.e. each ρ_g is an Automorphism and $\rho_{g_1} \circ \rho_{g_2} = \rho_{g_1 \circ g_2}$). G is called a **symmetry-group** if each ρ_g commutes with the one-parameter group $U(t)$:

$$[\rho_g, U(t)] = 0 \quad \text{for all } g \in G \tag{48}$$

where $U(t) = \exp(-\frac{i}{\hbar}Ht)$ describes the time-evolution of our Quantum-mechanical system. The reason for this definition becomes apparent when looking at it a bit closer: since G is a Lie Group we can write each element $g \in G$ as $g = \exp(sX)$ for some parameter $s \in \mathbb{R}$ and $X \in \mathrm{Lie}(G)$:

$$\begin{aligned} [\rho_{\exp(sX)}, U(t)] &= 0 \qquad \| \cdot \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} & (49)\\ \Longrightarrow \quad [\rho^*_X, U(t)] &= 0 & (50) \end{aligned}$$

This allows us now to show for any operator $A := \rho^*_X$:

$$\begin{aligned} W^A_{\Psi(t)} &= \langle U(t)\Psi, E^A U(t)\Psi\rangle & (51)\\ &\overset{[\rho_g, U(t)]=0}{=} \langle U(t)\Psi, U(t)E^A\Psi\rangle & (52)\\ &= \langle U(t)^* U(t)\Psi, E^A\Psi\rangle & (53)\\ &\overset{U(t)\ \text{unitary}}{=} \langle \Psi, E^A\Psi\rangle = W^A_\Psi & (54) \end{aligned}$$

which means that our condition for a symmetry entails the existence of an observable A for which the probability measure is conserved.

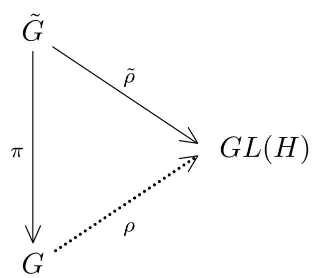
However, since any Automorphism is induced by a unitary map that is only unique safe a phase-factor, the symmetry of a group G does not necessarily entail the existence of a linear representation but only the existence of a something called a **projective representation** ρ with

$$\rho_{g_1} \circ \rho_{g_2} = \omega(g_1, g_2)\, \rho_{g_1 \circ g_2} \qquad \text{with } |\omega(g_1, g_2)| = 1 \tag{55}$$

a circumstance which is rather annoying since the theory we've developed so far for (Lie)groups and their representation is not applicable. It's therefore only natural to ask if there isn't some kind of a trick we could use to get rid of the phases ω and transform the projective representation ρ into something more manageable like a continuous, unitary representation for which we have already developed a lot of tools. The answer to that question is positive (with certain limitations[23]): for a large class of groups (especially for $SO(3)$) it is possible to choose a neighborhood G^0 of the unit element $e \in G$ such that $\rho|_{G^0}$ is a unitary and continuous representation. But if G is not simply connected (like e.g. $SO(3)$) than we cannot expect to be able to choose the phase-factor $\omega = 1$ globally. Instead we have to go from G to a 'larger' group $\tilde{G}$ which is simply connected and is called the universal covering group. G and $\tilde{G}$ are connected via a continuous, surjective map π:

$$\pi : G \longrightarrow \tilde{G} \tag{56}$$

which is even a differentiable group-homomorphism for the case that G is a connected Liegroup (like $SO(3)$). The great thing about the universal covering group is that the locally via π induced representation $\rho \circ \pi$ can be expanded *uniquely* to all of $\tilde{G}$ leaving us with the situation explained in the following diagram:



While $\tilde{\rho}$ and π are homomorphisms, ρ is only a projective representation (represented as a dotted line). In consequence the above diagram is only commutative modulo a phase factor λ

$$\tilde{\rho}(A) = \lambda \cdot \rho(\pi(A)) \tag{57}$$

with $|\lambda| = 1$. This means in particular that if $A \in \text{Kernel}\{\pi\}$, then $\tilde{\rho}(A) = \lambda \cdot \mathbb{1}$.

While all this may sound rather abstract and intangible, it has very real consequences for Quantum-mechanics. It allows us to "lift" the projective representations of $SO(3)$ to continuous, unitary representations of its universal covering group $SU(2)$ thereby introducing half integer spins.

2.4 $SU(2)$ - the universal covering group of $SO(3)$

From the previous paragraph it is apparent that in order to express rotational symmetry in Quantum mechanics we have to go from $SO(3)$ to its universal covering group. The

[23] see Straumann page 148

latter turns out to be $SU(2)$:

$$SU(2) := \{A \in U(2) : \det(A) = 1\} \tag{58}$$

where $U(2) := \{A \in GL(2,\mathbb{C}) : A^*A = \mathbb{1}\}$. While the proof for this assertion is beyond the scope of this script, it is important to at least know and understand how π connects $SO(3)$ and $SU(2)$. This shall be done in the following paragraphs step by step by 1) exploring $SU(2)$; 2) exploring its Lie-Algebra $su(2)$; 3) writing down π_U in a commutative diagram;

2.4.1 A few helpful facts about $SU(2)$

exercise: show for $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SU(2)$ the following: [24]

a) that the inverse exists and has the form $A^{-1} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}$;

b) that A can be simplified to $A = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$;

c) that there is a bijection between $SU(2)$ and the unit-sphere $S^2(\mathbb{C}^2)$ in $\mathbb{C}^2$;

2.4.2 A few helpful facts about $su(2)$

Lets start by quickly reiterating the definition of $su(2)$:

$$su(2) \quad = \quad \{X \in \mathrm{Mat}(2,\mathbb{C}) : \; X^* = -X \;\text{ and }\; \mathrm{tr}X = 0\} \tag{59}$$

This definition implies that any $X \in su(2)$ has the form $X = \begin{pmatrix} iz & x+iy \\ -x+iy & -iz \end{pmatrix}$ with $x, y, z \in \mathbb{R}$. This, of course, implies that $su(2)$ is a 3 dimensional vector-space. If we now define the following matrices:

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\mathbb{1}} \quad \underbrace{i\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\sigma_3} \quad \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{\sigma_2} \quad \underbrace{i\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{\sigma_1} \tag{60}$$

then it is again straight forward to see that the **Pauli Spin matrices** σ_1, σ_2 and σ_3 form a basis for $su(2)$.

---

[24] **ad a)** $\det(A) \neq 0 \Longrightarrow A^{-1}$ exists; for simple matrix multiplication shows that $A \circ A^{-1} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \circ \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} = (\alpha\delta - \gamma\beta)\mathbb{1}$. Since $\det(A) = \alpha\delta - \gamma\beta = 1$ it shows the assertion; **ad b)** Since the elements of $U(2)$ satisfy (by definition) $A^* = A^{-1}$ the expression for A is obvious with a); **ad c)** since the elements of $SU(2)$ satisfy (by definition) $\det(A) = 1$ the bijection is obvious;

2.4.3 The covering map π

We are now in a position to define π

$$\pi : SU(2) \longrightarrow SO(3) \tag{61}$$
$$A \longmapsto \pi_A \tag{62}$$

point-wise with the following cummutative diagram:

$$\begin{array}{ccc} \vec{x} \in \mathbb{R}^3 \longleftarrow \phi \longrightarrow & X := \vec{x} \circ \vec{\sigma} \in su(2) \\ \Big\downarrow \pi_A & \Big\downarrow \mathrm{Ad}_A(X) := AXA^{-1} \\ \vec{y} \in \mathbb{R}^3 \longleftarrow \phi \longrightarrow & Y =: \vec{y} \circ \vec{\sigma} \in su(2) \end{array}$$

Before we show now that π is indeed a homomorphism with Kernel$(\pi) = \{\pm \mathbb{1}\}$, it is necessary to have a closer look at the commutative diagram and show that is actually well defined.

Proof for π being well defined: In order to show that π is well defined, let's start by having a look at $\phi : \mathbb{R}^3 \longrightarrow su(2)$. Since the Pauli spin-matrices σ_i form a basis of $su(2)$, it is trivial to see that ϕ is an isomorphism, assigning each $X \in su(2)$ its coordinates $\vec{x} \in \mathbb{R}^3$ and vice versa. To see that ϕ is unitary show as an *exercise* that ...

i) ... $\langle X, Y \rangle := \frac{1}{2}\mathrm{tr}(X^* Y)$ is a scalar product on the $\mathbb{R}$ vector-space $su(2)$ i.e. that it is symmetric, bilinear and positive-definite[25];

ii) ... ϕ conserves the scalar product i.e. that $\langle \vec{x}, \vec{y} \rangle = \langle \phi(\vec{x}), \phi(\vec{y}) \rangle$[26];

In order to show that for any $A \in SU(2)$ the map $\mathrm{Ad}_A : su(2) \longrightarrow su(2)$ is indeed an orthogonal endomorphism we need to work a bit harder and verify ...

i) ... that Ad_A is indeed an endomorphism (i.e. that $\mathrm{Ad}_A(X) \equiv AXA^{-1} \in su(2)$) by showing that AXA^{-1} is traceless and anti-hermitian:
traceless: $\mathrm{tr}(AXA^{-1}) \overset{\text{cyclic}}{=} \mathrm{tr}(A^{-1}AX) = \mathrm{tr}(X) \overset{X \in su(2)}{=} 0$;

[25] **symmetric:** follows straight from trace being cyclic (cyc.) and X, Y being anti-hermitian (a.-h.): $\langle X, Y \rangle \equiv \frac{1}{2}\mathrm{tr}(X^*Y) \overset{\text{cyc.}}{=} \frac{1}{2}\mathrm{tr}(YX^*) \overset{\text{a.-h.}}{=} \frac{1}{2}\mathrm{tr}((-Y^*)(-X)) = \langle Y, X \rangle$; **bilinear:** follows straight forward from trace being linear i.e. $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ and any coefficients $\mu \in \mathbb{R}$; **positive definite:** $\langle X, X \rangle \overset{\text{Def}}{=} \frac{1}{2}\mathrm{tr}(X^*X) \overset{\text{Def.}}{=} \sum_i \sum_k (X^*_{ik} X_{ki}) = \sum_i \sum_k (\overline{X}_{ki} X_{ki}) = \sum_{i,k} |X_{ki}|^2 \geq 0$ and $= 0 \iff X = 0$;

[26] in order to show that ϕ conserves the scalar product we only need to show it for the image of an orthonormal base such as $\{\mathbf{e}_i\}$ and $\{\phi(\mathbf{e}_i) = \sigma_i\}$; since $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \langle \sigma_i, \sigma_j \rangle = \delta_{ij}$ we are done;

anti hermitian: **a)** preliminary step: $(AB)^* = B^*A^*$ because $((AB)^*)_{ij} = (\overline{AB}^T)_{ij} = (\overline{AB})_{ji} = \overline{(AB)_{ji}} = \sum_k \overline{A}_{jk}\overline{B}_{ki} = \sum_k A^*_{kj}B^*_{ik} = (B^*A^*)_{ij}$; **b)** because of a) we can now see that $(ABC)^* = C^*B^*A^*$ because $(ABC)^* = ((AB)C)^* = C^*(AB)^* = C^*B^*A^*$; **c)** because of b) and since by definition $A^* = A^{-1}$ and $X^* = -X$ we can finally conclude that $(AXA^{-1})^* = A^{-1^*}X^*A^* = A^{**}(-X)A^* = -(AXA^*)$

ii) ... that $\mathrm{Ad}_A(\cdot)$ is linear and orthogonal:
linear: $\mathrm{Ad}_A(X+Y) \equiv A(X+Y)A^{-1} = AXA^{-1} + AYA^{-1} \equiv \mathrm{Ad}_A(X) + \mathrm{Ad}_A(Y)$
orthogonal:

$$\begin{aligned}
\langle \mathrm{Ad}_A(X), \mathrm{Ad}_A(Y)\rangle &\equiv \mathrm{tr}((AXA^{-1})^*AYA^{-1}) \quad \text{by definition} \\
&= \mathrm{tr}((A^{-1})^*X^*A^*AYA^{-1}) \quad \text{since}(ABC)^* = C^*B^*A^* \\
&= \mathrm{tr}(\underbrace{A^{-1}(A^{-1})^*}_{\mathbb{1}}X^*\underbrace{A^*A}_{\mathbb{1}}Y) \quad \text{using tr}(\cdot)\text{ cyclic and } A^* = A^{-1} \\
&= \mathrm{tr}(X^*Y) = \langle X, Y\rangle
\end{aligned}$$

Proof for π being the universal covering group:
Let's start by showing that π is linear (i.e. that $\pi_{AB} = \pi_A \circ \pi_B$). With the commutative diagram shown at the beginning of this subsection it is easy to see that π is linear if and only if $\mathrm{Ad} : SU(2) \longrightarrow \mathrm{End}(su(2))$ is linear which is the case as can be seen in the following:

$$\mathrm{Ad}_{AB}(X) \equiv (AB)X(AB)^{-1} = A(BXB^{-1})A^{-1} = \mathrm{Ad}_A(\mathrm{Ad}_B(X)) \tag{63}$$

Having shown that π is indeed linear, we would now like to show that $\mathrm{Kernel}(\pi) = \{\pm\mathbb{1}\}$. Again the commutative diagram becomes helpful as it is apparent that $\mathrm{Kernel}(\pi) = \mathrm{Kernel}\{\mathrm{Ad}\}$. Therefore we are looking for all matrices $A \in SU(2)$ for which $\mathrm{Ad}_A = \mathrm{id}$ and consequently for all $A \in SU(2)$ for which $AXA^{-1} = X$ for any $X \in su(2)$. Obviously this is the case for $A = \pm\mathbb{1}$. The question is now if $A = \pm\mathbb{1}$ is not only a sufficient but also a necessary condition. Let us therefore consider

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{X\in su(2)} = \underbrace{\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}}_{A\in SU(2)} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{X} \cdot \underbrace{\begin{pmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{pmatrix}}_{A^*=A^{-1}} \tag{64}$$

$$= \underbrace{\begin{pmatrix} |\alpha|^2 - |\beta|^2 & -2\alpha\beta \\ -2\overline{\alpha}\overline{\beta} & |\beta|^2 - |\alpha|^2 \end{pmatrix}}_{AXA^{-1}} \tag{65}$$

$$\Longrightarrow \quad |\alpha|^2 - |\beta|^2 = 1 \quad \text{and} \quad \alpha\beta = 0 \tag{66}$$

$$\Longrightarrow \quad \beta = 0 \quad \text{and} \quad |\alpha|^2 = 1 \tag{67}$$

This last line shows that $A = \pm\mathbb{1}$ is not only sufficient but also necessary for $\mathrm{Ad}_A = \mathrm{id}$ and therefore $\mathrm{Kernel}(\pi) = \{\pm\mathbb{1}\}$.

The fact that π is surjective shall only be scetched here:

- first show that for $A = \begin{pmatrix} \cos\theta/2 & -i\sin\theta/2 \\ -i\sin\theta/2 & \cos\theta/2 \end{pmatrix} \in SU(2)$ the map $\mathrm{Ad}_A(\cdot) : su(2) \longrightarrow su(2)$ can be expressed in the basis $\{\sigma_i\}_{i=1,2,3}$ as:

$$\mathrm{Ad}_A(\cdot)|_{\sigma_i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} = R(\mathbf{e}_1, \theta) \tag{68}$$

- next show that for $B = \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{pmatrix} \in SU(2)$ the map $\mathrm{Ad}_A(\cdot) : su(2) \longrightarrow su(2)$ can be expressed in the basis $\{\sigma_i\}_{i=1,2,3}$ as:

$$\mathrm{Ad}_B(\cdot)|_{\sigma_i} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} = R(\mathbf{e}_3, \phi) \tag{69}$$

- finally we can use a theorem from linear Algebra (Felder page 45) that any $R \in SO(3)$ can be expressed as:

$$R = R(\mathbf{e}_3, \phi) R(\mathbf{e}_1, \theta) R(\mathbf{e}_3, \psi) \tag{70}$$

2.5 $SU(2)$ and the angular momentum operator J_k

In the previous section we saw for a simple example (namely for a QM system described in a Hilbertspace $\mathcal{H} = L^2(\mathbb{R}^3)$) that if we define the angular momentum operator L_k through the representation ρ^* induced by a rotation as

$$L_k := i\hbar\rho^*(I_k) \tag{71}$$

then L_k coincides with the operator one would obtain by simply using the correspondence principle. This is already a strong indication that such a definition through the representation on the Lie-Algebra is sensible. With the tools obtained in this chapter we would like to give another reason for this definition:

Let us consider a rotation invariant QM system (such as an electron, atom, nucleus and the like) described by a ray in a separable Hilbertspace $\mathcal{H}$. We know now that the symmetry of this system is described by a projective representation $\rho : SO(3) \longrightarrow GL(\mathcal{H})$ that can be "lifted" to a continuous, unitary representation $\tilde{\rho} : SU(2) \longrightarrow GL(\mathcal{H})$. Because it is a symmetry, we know that ρ commutes with the time evolution operator $U(t)$

$$[\tilde{\rho}(A), U(t)] = 0 \text{ for all } t \in \mathbb{R} \text{ and } A \in SU(2) \tag{72}$$

where $U(t) := \exp -\frac{i}{\hbar}Ht$. If we now generalize our considerations from before and define the angular momentum operator as

$$J_k := i\hbar\ \tilde{\rho}^*(I_k) \tag{73}$$

then it follows immediately from equ. 72 that $[J_k, U(t)] = 0$ and hence that J_k is an *integral of motion*[27]. This is exactly what we would expect from classical mechanics and a further indication that our Definition in eq. 73 is sound.

[27] see Straumann page 129 and Wikipedia

3 Tensorproductspaces for elementary particles and their angular momentum

As it has been mentioned already earlier, the Stern-Gerlach experiment suggests that the electron carries some kind of an intrinsic angular momentum which cannot be explained by describing it as a ray $[\Psi] \in L^2(\mathbb{R}^3)$. We therefore have to find a more adequate Hilbertspace to describe our system like

$$\Psi(\mathbf{x}) = \begin{pmatrix} \Psi_1(\mathbf{x}) \\ \vdots \\ \Psi_n(\mathbf{x}) \end{pmatrix} \tag{74}$$

This “guess” is not only straight forward but also quite natural considering the fact that classical fields are not described as scalar functions neither (Straumann page 66 bottom). For practical reasons that will become apparent shortly, it is useful to make the following identification of Vectorspaces (via the canonical Isomorphism):

$$\underbrace{L^2(\mathbb{R}^3) \times \ldots \times L^2(\mathbb{R}^3)}_{\text{n times}} \simeq L^2(\mathbb{R}^3) \otimes \mathbb{C}^n \tag{75}$$

Once we understand these tensor-product spaces and how representations operate on them, it will be straight forward to define our angular momentum operator J_k for such systems explicitly.

3.1 Tensorproducts of Vectorspaces

The tensor product $\mathcal{U} \otimes \mathcal{V}$ of two $\mathbb{K}$-vector spaces $\mathcal{U}$ and $\mathcal{V}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$) can be introduced in different ways that vary in complexity. Here I would like to follow a simple approach (taken from Felder page 87) that defines the tensor product as a quotient set (the set of equivalence classes) on the linear envelope $\mathcal{L}$ of the Cartesian product $\mathcal{U} \times \mathcal{V}$.

Let $\mathcal{L}(\mathcal{U} \times \mathcal{V})$ be the set of all finite linear combinations of vector-pairs $(u_i, v_i) \in \mathcal{U} \times \mathcal{V}$:

$$\mathcal{L}(\mathcal{U} \times \mathcal{V}) = \left\{ \sum_{i=1}^{n} \alpha_i (u_i, v_i) : \quad n \in \mathbb{N} \text{ and } \alpha_i \in \mathbb{K} \right\} \tag{76}$$

On this set an equivalence relation $\overset{R}{\sim}$ shall be defined:

(i) $(u_1 + u_2, v) \sim (u_1, v) + (u_2, v)$

(ii) $(u, v_1 + v_2) \sim (u, v_1) + (u, v_2)$

(iii) $\lambda(u, v) \sim (\lambda u, v) \sim (u, \lambda v)$

exercise: Show that $\overset{R}{\sim}$ is indeed an equivalence relation.

The equivalence classes $[(v,w)]$ are commonly denoted as $v \otimes w$ and the set of all equivalence classes is the tensor-product

$$\mathcal{U} \otimes \mathcal{V} \equiv \mathcal{L}(\mathcal{U} \times \mathcal{V}) / \overset{R}{\sim} \tag{77}$$

which has the following properties:

(i) All elements of $\mathcal{U} \otimes \mathcal{V}$ are either pure tensors $u \otimes v$ or linear combinations of pure tensors $u_1 \otimes v_1 + ... + u_n \otimes v_n$

(ii) Let $\{\mu_i\}_{i \in I}$ and $\{\nu_j\}_{j \in J}$ be a basis in $\mathcal{U}$ and $\mathcal{V}$ respectively. Then $\{\mu_i \otimes \nu_j\}_{(i,j) \in I \times J}$ is a basis in $\mathcal{U} \otimes \mathcal{V}$

(iii) The definition of the equivalence relation $\overset{R}{\sim}$ entails some simple calculation rules:

- $(u_1 + u_2) \otimes v = u_1 \otimes v \; + u_2 \otimes v$
- $u \otimes (v_1 + v_2) = \; u \otimes v_1 + \; u \otimes v_2$
- $\lambda(u \otimes v) = (\lambda u) \otimes v = u \otimes (\lambda v)$ where λ is an element of the field $\mathbb{K}$

The proof for these properties can be found in Felder page 87.

If we now consider two operators $A \in \mathrm{Hom}(\mathcal{U})$ and $B \in \mathrm{Hom}(\mathcal{V})$ then the tensor-product $A \otimes B$ on $\mathcal{U} \otimes \mathcal{V}$ is defined canonically as:

$$(A \otimes B)(u \otimes v) := (Au) \otimes (Bv) \tag{78}$$

To make all this a bit more tangible and in order to gain some experience with these new definitions, let us consider the simple example in which $\mathcal{U} = \mathbb{R}^3$ with Basis $\mathcal{B}_{\mathcal{U}} = \{e_1, e_2, e_3\}$ and $\mathcal{V} = \mathbb{R}^2$ with basis $\mathcal{B}_{\mathcal{V}} = \{f_1, f_2\}$. For $\mathcal{U} \otimes \mathcal{V}$ basis' $\mathcal{B}_{\mathcal{U} \otimes \mathcal{V}}$ we have in theory $6! = 720$ options but 2 of them are most natural:

- option 1: $\mathcal{B}_{\mathcal{U} \otimes \mathcal{V}} = \{e_1 \otimes f_1, e_2 \otimes f_1, e_3 \otimes f_1, e_1 \otimes f_2, e_2 \otimes f_2, e_3 \otimes f_2\}$
- option 2: $\mathcal{B}_{\mathcal{U} \otimes \mathcal{V}} = \{e_1 \otimes f_1, e_1 \otimes f_2, e_2 \otimes f_1, e_2 \otimes f_2, e_3 \otimes f_1, e_3 \otimes f_2\}$

Choosing option 1, it is straight forward to see now that $A \otimes \mathrm{id}$ and $\mathrm{id} \otimes B$ have the form:

$$A \otimes \mathrm{id} = \left(\begin{array}{c|c} A & 0 \\ \hline 0 & A \end{array}\right) \qquad \text{while} \qquad \mathrm{id} \otimes B = \left(\begin{array}{c|c} b_{11}\mathbb{1} & b_{12}\mathbb{1} \\ \hline b_{21}\mathbb{1} & b_{22}\mathbb{1} \end{array}\right) \tag{79}$$

exercise By applying $A \otimes B$ on each element of the basis $\mathcal{B}_{\mathcal{U}\otimes\mathcal{V}}$ and writing out the result in Matrix notation, show that

$$A \otimes B \;=\; A \otimes \mathrm{id} \;\cdot\; \mathrm{id} \otimes B \tag{80}$$

3.2 Representations on Tensorproducts

Having introduced the tensor-product space $\mathcal{U}\otimes\mathcal{V}$ it is now important to understand, how a representation and its derivative work on such a space. To this end let us assume that the representations $\rho^{\mathcal{U}}$ and $\rho^{\mathcal{V}}$ are given

$$\rho^{\mathcal{U}}: \quad G \longrightarrow GL(\mathcal{U}) \tag{81}$$
$$\rho^{\mathcal{V}}: \quad G \longrightarrow GL(\mathcal{V}) \tag{82}$$

We can now sensibly define the tensorproduct of those two representations as follows:

$$\rho^{\mathcal{U}} \otimes \rho^{\mathcal{V}}: \quad G \times G \longrightarrow GL(\mathcal{U}\otimes\mathcal{V}) \tag{83}$$
$$(g,\tilde{g}) \mapsto (\rho^{\mathcal{U}} \otimes \rho^{\mathcal{V}})_{(g,\tilde{g})} := \rho^{\mathcal{U}}_g \otimes \rho^{\mathcal{V}}_{\tilde{g}} \tag{84}$$

This will become very useful for describing rotations in Quantum Mechanics and allow us to define the angular momentum operator as $J_k = i\hbar(\rho^{\mathcal{U}} \otimes \rho^{\mathcal{V}})^*$. Before we do the calculation, let us remind ourselves of the commutative diagram

$$\begin{array}{ccc} G \times G & \xrightarrow{\rho^{\mathcal{U}}\otimes\rho^{\mathcal{V}}} & \mathrm{GL}(\mathcal{U}\otimes\mathcal{V}) \\ {\scriptstyle (\exp(t\cdot),\exp(t\cdot))}\uparrow & & \uparrow{\scriptstyle (\exp(t\cdot),\exp(t\cdot))} \\ \mathcal{G}\oplus\mathcal{G} & \xrightarrow[(\rho^{\mathcal{U}}\otimes\rho^{\mathcal{V}})^*]{} & \mathrm{gl}(\mathcal{U}\otimes\mathcal{V}) \end{array} \tag{85}$$

which shows how $(\rho^{\mathcal{U}} \otimes \rho^{\mathcal{V}})^*$ is operating (where $\mathcal{G}$ denotes the Liegroup of G).

$$(\rho^{\mathcal{U}} \otimes \rho^{\mathcal{V}})^*_{X,Y} \quad := \quad \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(\rho^{\mathcal{U}}_{\exp(tX)} \otimes \rho^{\mathcal{V}}_{\exp(tY)}\right) \tag{86}$$
$$\overset{\text{eq. 80}}{=} \quad \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(\rho^{\mathcal{U}}_{\exp(tX)} \otimes \mathbb{1} \;\cdot\; \mathbb{1} \otimes \rho^{\mathcal{V}}_{\exp(tY)}\right) \tag{87}$$
$$= \quad \left(\rho^{\mathcal{U}}\right)^*_X \otimes \mathbb{1} \;+\; \mathbb{1} \otimes \left(\rho^{\mathcal{V}}\right)^*_Y \tag{88}$$

It is important to note at this point that the above calculation is still valid if we set $Y = X$. For that case we are only looking at the diagonal elements $(g,g) \in G \times G$ (which is the way that G can be embedded in $G \times G$) and write the tensor product as $\otimes_{\mathrm{i}}$:

$$\rho^{\mathcal{U}} \otimes_{\mathrm{i}} \rho^{\mathcal{V}}: \quad G \longrightarrow GL(\mathcal{U}\otimes\mathcal{V}) \tag{89}$$
$$g \mapsto (\rho^{\mathcal{U}} \otimes_{\mathrm{i}} \rho^{\mathcal{V}})_{(g,\tilde{g})} := \rho^{\mathcal{U}}_g \otimes \rho^{\mathcal{V}}_g \tag{90}$$

3.3 The angular Momentum and Spin Operator

With these tools we now only have to make a good "guess" on the Hilbertspace $\mathcal{H}$ the electron will be living in, in oder to define its angular momentum operator J_k. So far we only argued that $\mathcal{H}$ will have the form $\mathcal{H} = L^2(\mathbb{R}^3) \otimes \mathbb{C}^n$ without specifying n.

This good guess turns out to be

$$\mathcal{H} = L^2(\mathbb{R}^3) \otimes \mathbb{C}^2 \tag{91}$$

and there are good reasons to choose $n = 2$ that are connected to the irreducible representations of $SU(2)$ which however we haven't studied so far. So for now I have to ask the reader to take this choice in good faith assuring him or her that it will become obvious once we study $SU(2)$ in more detail. There is, however, an experimental fact that is strongly related to the choice of $n = 2$ and that is the fact that the beam under observation in the Stern-Gerlach experiment actually splits up into 2 beams when going through an inhomogeneous magnetic field.

Let us now consider an electron in a system with rotational symmetry. As we learned before such a symmetry can be expressed as a projective representation of $SO(3)$

$$\rho : SO(3) \longrightarrow GL(L^2(\mathbb{R}^3) \otimes \mathbb{C}^2) \tag{92}$$

$$R \longmapsto \rho_R \quad \text{with} \quad \rho_{R_1} \circ \rho_{R_2} = \omega(R_1, R_2)\, \rho_{R_1 \circ R_2} \tag{93}$$

where in general we only know that $|\omega(g_1, g_2)| = 1$. Such a map is given by

$$(\rho_R \Psi)(x) := \psi(R^{-1}x) \otimes \mathcal{S}_R\, \xi \quad \text{with} \quad \mathcal{S}_{R_1} \circ \mathcal{S}_{R_2} = \pm \mathcal{S}_{R_1 R_2} \tag{94}$$

As with the choice for $\mathcal{H}$, this is again only an educated "guess" which can only be justified later, when we actually see the results obtained with it. For the moment I would just like to point out that this choice ...

(a) ... has physical merit because it incorporates $\Psi \longmapsto \Psi(R^{-1})$

(b) ... is the simplest choice for a projective representation that is non-trivial (i.e. where the phase-factor ω is not just simply chosen to be 1).

As has been discussed previously it is now possible to "lift" ρ to a unitary representation $\tilde{\rho}$ of $SU(2)$ which shall be denoted *symbolically* as:

$$\tilde{\rho}_U \;:=\; \psi(\pi(U)^{-1}\cdot) \otimes \tilde{\mathcal{S}}_U \quad \text{with } U \in SU(2) \tag{95}$$

With this representation $\tilde{\rho}$ we can now define the total angular momentum operator J_k as

$$J_k \;:=\; i\hbar\, \tilde{\rho}^*_{\sigma_k} \tag{96}$$

$$\overset{\text{eq. 88}}{=} \; L_k \otimes \mathbb{1} \;+\; \mathbb{1} \otimes S_k \tag{97}$$

where L_k is defined as in eq. 35 and refers to the electron's orbital angular momentum while $S_i := i\hbar\, \tilde{\mathcal{S}}^*_{I_i}$ is interpreted as the electron's intrinsic angular momentum.

4 SU(2) and its irreducible representations

In this chapter we will study the irreducibel representation of $SU(2)$. This can be done using either the global or the infinitessimal method. Since the former requires a mathematical tool we haven't discussed so far (the "Haar-measure") I will focus on the latter and only outline the main results of the global method leaving it up to the individual to read up on it in further detail. At the end of this chapter the following results will hopefully be more clear:

- The irreducible representations D^j of $SU(2)$ can be indexed with $j = 0, \frac{1}{2}, 1, \frac{3}{2}, ...$ and are therefore countable. Their dimension is given by:
$$\dim(D^j) = 2j + 1 \tag{98}$$
- The angular momentum operator $L^2 := \sum_{k=1,2,3} {L_k}^2$ induced by this representation takes on the simple form
$$L^2 = j(j+1)\mathrm{id} \tag{99}$$
where $L_k := \ i\hbar\,\{D^j\}^*(\sigma_k)$ is defined on the Lie Algebra level.
- The character χ^j of the irreducible representation D^j is given by $\chi^j(\lambda) = \sum_{k=-j}^{+j} \lambda^{2k}$ leading us to the well known **Clebsch-Gordan-Series** for adding angular momentum in Quantum mechanics:
$$D^{j_1} \otimes_{\mathrm{i}} D^{j_2} = D^{|j_1-j_2|} \oplus ... \oplus D^{|j_1+j_2|} \tag{100}$$
- The value for the irreducible representation D^j is given by:
$$D^j(-\mathbb{1}) = (-1)^{2j}\mathbb{1} \tag{101}$$
This entails that all even numbered representations D^j on $SU(2)$ induce a unique representation on $SO(3)$ while the odd numbered induce an ambiguous representation on $SO(3)$.

4.1 The infinitesimal method

Let us consider an irreducible representation $D : SU(2) \longrightarrow GL(\mathcal{V})$ on a finite-dimensional vector-space $\mathcal{V}$ with $\dim\mathcal{V} = n + 1$ with $n \in \mathbb{N}_0$. Since $SU(2)$ is simply connected, D is irreducible if and only if $D^* : su(2) \longrightarrow gl(\mathcal{V})$ is irreducible which allows us to derive the properties of D mentioned in the introduction on the level of D^*. To this end let us define the operators
$$L_i := iD^*(\sigma_i) \qquad \text{and} \qquad L_\pm := L_1 \pm iL_2 \tag{102}$$

(where $i = 1, 2, 3$ and $\hbar$ was omitted for simplicity reasons) which have the following commutator relationships:

$$
\begin{aligned}
[L_i, L_j] &= i\epsilon_{ijk} L_k \quad \text{as shown in eq. 36} & (103)\\
[L_3, L_\pm] &= [L_3, L_1] \pm [L_3, iL_2] = iL_2 \pm (-i^2) L_1 = \pm L_\pm & (104)\\
[L_+, L_-] &= [L_1 + iL_2, L_1 - iL_2] = -2i[L_1, L_2] = 2L_3 & (105)
\end{aligned}
$$

Step 1 - constructing a basis in $\mathcal{V}$: From linear Algebra we know that for L_3 linear we have at least one complex Eigen-value $\tilde{\lambda}$ with Eigen-vector $v_{\tilde{\lambda}} \neq 0$. With the above equations it is furthermore easy to see that:

$$L_3(L_\pm v_{\tilde{\lambda}}) = L_\pm(L_3 v_{\tilde{\lambda}}) + [L_3, L_\pm] v_{\tilde{\lambda}} \stackrel{\text{eq. 105}}{=} (\tilde{\lambda} \pm 1) L_\pm v_{\tilde{\lambda}} \qquad (106)$$

This shows that $L_\pm v_{\tilde{\lambda}}$ is an Eigen-vector for L_3 with Eigen-value $\tilde{\lambda} \pm 1$. This allows us now to find a basis of $\mathcal{V}$ by iteratively applying $L_\pm$ to $v_{\tilde{\lambda}}$. But since $\mathcal{V}$ is finite dimensional it is obvious that for some $\tilde{n} \in \mathbb{N}$ this iterative chain must end:

$$L_3(L_+^{\tilde{n}} v_{\tilde{\lambda}}) = \underbrace{(\tilde{\lambda} \;+\; \tilde{n})}_{\lambda} \underbrace{L_+^{\tilde{n}} v_{\tilde{\lambda}}}_{v_\lambda} \qquad \text{and} \qquad L_+^{\tilde{n}+1} v_{\tilde{\lambda}} = 0 \qquad (107)$$

Now we can go the opposite direction by applying L_- to v_λ iteratively and will again encounter a termination-condition:

$$L_3(L_-^n v_\lambda) = (\lambda \;-\; n) L_-^n v_\lambda \qquad \text{and} \qquad L_-^{n+1} v_\lambda = 0 \qquad (108)$$

While the number of steps $\tilde{n}$ going "up" to the termination-condition were unknown, the number of steps coming "down" is equal to the dimension n of the vector-space $\mathcal{V}$. This shall be illustrated in the following sketch:

$$
\begin{array}{ccccccccc}
 & & v_{\tilde{\lambda}} & \overbrace{\xrightarrow{L_+} \; v_{\tilde{\lambda}+1} \; \xrightarrow{L_+} \; \ldots \; \xrightarrow{L_+}}^{\tilde{n}-\text{times}} & v_{\tilde{\lambda}+\tilde{n}} & \xrightarrow{L_+} 0 \\
 & & & & \| & \\
0 \xleftarrow{L_-} & v_{\lambda-n} & & \underbrace{\xleftarrow{L_-} \; \ldots \; \xleftarrow{L_-} \; v_{\lambda-1} \; \xleftarrow{L_-}}_{n-\text{times}} & v_\lambda &
\end{array} \qquad (109)
$$

Step 2 - finding a relationship between λ and n: From a physics point of view λ is an interesting variable because it is the maximum value that can be observed for the z-component of the angular momentum. It would therefore be desirable to index the irreducible representations with λ rather than their dimension n. That is the motivation for us to derive an analytic relation between those two variables in the following:

$$L_+ v_{\lambda-1} \overset{\text{Def.}}{=} L_+(L_- v_\lambda) = L_-(\underbrace{L_+ v_\lambda}_{\overset{\text{eq. }109}{=} 0}) + \underbrace{[L_+, L_-]}_{\overset{\text{eq. }105}{=} 2L_3} v_\lambda \overset{\text{eq. }108}{=} \underbrace{2\lambda}_{=:\mu_{\lambda-1}} v_\lambda$$

$$L_+ v_{\lambda-2} \overset{\text{Def.}}{=} L_+(L_- v_{\lambda-1}) = L_-(\underbrace{L_+ v_{\lambda-1}}_{\mu_{\lambda-1} v_\lambda}) + \underbrace{[L_+, L_-]}_{\overset{\text{eq. }105}{=} 2L_3} v_{\lambda-1} \overset{\text{eq. }108}{=} \underbrace{(\mu_{\lambda-1} + 2(\lambda-1))}_{=:\mu_{\lambda-2}} v_{\lambda-1}$$

It is now easy to see the more general expression:

$$L_+ v_{p-1} = \underbrace{\mu_p + 2p}_{\mu_{p-1}} v_p \quad \text{for } p = \lambda, \lambda-1, \lambda-2, ..., \lambda-(n-1) \quad \text{and} \quad \mu_\lambda \equiv 0 \tag{110}$$

By inserting our previous results we can then arrive at an explicit expression for μ_{p-1}

$$\mu_{p-1} = 2\lambda + 2(\lambda-1) + ... + 2p \tag{111}$$

$$= 2\underbrace{(\lambda + (\lambda-1) + ... + p)}_{\lambda-p+1 \text{ terms}} \tag{112}$$

$$= 2\left[\frac{1}{2}(\lambda+p)(\lambda-p+1)\right] = (\lambda+p)(\lambda-p+1) \tag{113}$$

We now use again a termination condition to derive a direct relation between n and λ

$$L_+ \underbrace{(L_- v_{\lambda-n})}_{0} = L_-(\underbrace{L_+ v_{\lambda-n}}_{\mu_{\lambda-n} v_{\lambda-n+1}}) + \underbrace{[L_+, L_-]}_{\overset{\text{eq. }105}{=} 2L_3} v_{\lambda-n} \tag{114}$$

$$\overset{\text{eq. }108}{=} [\mu_{\lambda-n} + 2(\lambda-n)] \underbrace{v_{\lambda-n}}_{\neq 0} \tag{115}$$

$$\implies 0 = \underbrace{[\mu_{\lambda-n} + 2(\lambda-n)]}_{=:\mu_{\lambda-n-1}} \tag{116}$$

$$\overset{\text{eq. }113}{=} [\lambda + (\underbrace{\lambda-n}_{p})][\lambda - (\underbrace{\lambda-n}_{p}) + 1] \tag{117}$$

$$= (2\lambda-n)(n+1) \tag{118}$$

$$\implies \boxed{2\lambda = n} \tag{119}$$

It is therefore legitimate to label the irreducible representations with an index[28]:

$$\boxed{\begin{array}{rcl} \lambda &=& \frac{n}{2} \quad \text{with } n \in \mathbb{N}_0 \\ \dim(D^\lambda) &=& 2\lambda + 1 \end{array}} \tag{120}$$

With all these preparation it is now fairly easy to derive the eigenvalues of $L^2 = \sum_i L_i^2$ which we already used in section 1.4. This will be done again in two steps: in step 1 the relationship $L^2 = L_3^2 + L_3 + L_- L_+$ will be shown and in step 2 we are going to show that $L^2 v_{p-1} = \lambda(\lambda + 1)$.

Step 1 - deriving the expression $L^2 = L_3^2 + L_3 + L_- L_+$: Let's start by having a look at the last two terms:

$$\begin{aligned} L_3 + L_- L_+ &\overset{\text{eq. 105}}{=} \frac{1}{2}[L_+, L_-] + L_- L_+ && (121)\\ &= \frac{1}{2}(L_+ L_- + L_- L_+) && (122)\\ &\overset{\text{Def.}}{=} \frac{1}{2}\left\{(L_1 + iL_2)(L_1 - iL_2) + (L_1 - iL_2)(L_1 + iL_2)\right\} && (123)\\ &= L_1^2 \ + \ L_2^2 && (124) \end{aligned}$$

With that it is easy to see the assertion.

Step 2 - deriving the Eigen-value $\lambda(\lambda+1)$ of L^2: with the expression $L^2 = L_3^2 + L_3 + L_- L_+$ derived in step 1, it is easy to see that:

$$L^2 v_\lambda = \left\{L_3^2 + L_3 + L_- L_+\right\} v_\lambda \overset{L_+ v_\lambda = 0}{=} (\lambda^2 + \lambda) v_\lambda = \lambda(\lambda + 1) v_\lambda \tag{125}$$

which proofs the assertion. For $p = \lambda, \lambda - 1, \lambda - 2, ..., \lambda - (n - 1)$ we have to work a bit harder:

$$\begin{aligned} L^2 v_{p-1} &= \left\{L_3^2 + L_3 + L_- L_+\right\} v_{p-1} && (126)\\ &\overset{\text{eq. 113}}{=} \left\{(p-1)^2 + (p-1) + (\lambda + p)(\lambda - p + 1)\right\} v_{p-1} && (127)\\ &= \left\{(p^2 - 2p + 1) + (p - 1) + (\lambda^2 - \lambda p + \lambda + p\lambda - p^2 + p)\right\} v_{p-1} && (128)\\ &= \lambda(\lambda + 1) v_{p-1} && (129) \end{aligned}$$

This shows that any vector in $\mathcal{V}$ has the same L^2-Eigenvalue $\lambda(\lambda + 1)$ that determined by the irreducible representation D^λ associated with $\mathcal{V}$.

[28] strictly speaking we would still have to proof that the irreducible representation we found here are really *all* there are; For further details you can consult Straumann page 153.

4.2 The global method

As mentioned at the beginning of this chapter the global method requires the knowledge of the so called "Haar-measure", a mathematical tool we haven't acquired so far. Therefore I will only outline its general idea in this section. For more details please consult Straumann, page 153.

Let us consider the vector-space $\mathcal{V}$ of all polynomials $\Psi^j_m(\xi_1,\xi_2)$ in 2 complex variables of degree $2j$ with $j = 0, \frac{1}{2}, 1, ...$:

$$\mathcal{V}^j := \{\sum_{m=-j}^{+j} \lambda_m \Psi^j_m(\xi_1,\xi_2): \quad \lambda_m, \xi_1, \xi_2 \in \mathbb{C}\} \tag{130}$$

$$\text{where} \quad \Psi^j_m(\xi_1,\xi_2) := \frac{1}{\sqrt{(j+m)!(j-m)!}} \xi_1^{j+m}\xi_2^{j-m} \tag{131}$$

On this vector-space it is straight forward to define a representation D^j:

$$D^j: \quad SU(2) \longrightarrow GL(\mathcal{V}^j) \tag{132}$$

$$U \longmapsto D^j_U \quad \text{with} \quad D^j_U \Psi(\xi) := \Psi(U^T\xi) \tag{133}$$

Of course it needs to be verified that this is indeed a representation. But if we assume for now that it is, and if we furthermore assume to know that each and every conjugacy classes of $SU(2)$ can be represented with

$$U_\lambda := \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} \quad \text{with} \quad |\lambda| = 1 \quad \text{and } [U_\lambda] = [U_{\bar{\lambda}}] \tag{134}$$

then it is straight forward to derive an analytic expression for the character $\chi^j(U_\lambda) \equiv \text{tr}(D^j_{U_\lambda})$ of D^j by looking at the image $D^j_{U_\lambda}\Psi^j_m(\xi)$ of the basis-vectors:

$$D^j_{U_\lambda}\Psi^j_m(\xi) \overset{\text{Def.}}{=} \Psi^j_m(U_\lambda^T\xi) = \Psi^j_m(\lambda\xi_1, \bar{\lambda}\xi_2) \tag{135}$$

$$\overset{\text{Def.}}{=} \lambda^{j+m}\bar{\lambda}^{j-m}\Psi^j_m = \lambda^{j+m}\lambda^{-(j-m)}\Psi^j_m \tag{136}$$

$$= \lambda^{2m}\Psi^j_m \tag{137}$$

With this result it is straightforward to see that:

$$\boxed{\begin{array}{rcl} D^j(-\mathbb{1}) & = & (-1)^{2j}\mathbb{1} \\ \chi^j(U_\lambda) & = & \sum_{m=-j}^{+j}\lambda^{2m} \end{array}} \tag{138}$$

At this point it would be easy to proof that D^j is indeed irreducible by showing that

$$\langle \chi^j, \chi^j \rangle \;=\; 1 \tag{139}$$

if it wasn't for the fact that $SU(2)$ is an infinite group. For the latter we haven't discussed how to sensibly define the scalar-product so far and it is beyond the scope of these notes to do so here. But we can mention in bypassing that it would require the definition of the so called "Haar-measure" which can be found again in Straumann (Appendix B) and allows one to show that the previous equation is indeed true.

Knowing the character of the irreducible representations it is now easy to show that:

$$D^{j_1} \otimes_{\mathrm{i}} D^{j_2} = D^{|j_1-j_2|} \oplus ... \oplus D^{|j_1+j_2|} \tag{140}$$

which is known as the **Clebsch-Gordan series**. Since we know from eq. 80 that:

$$A \otimes B = \left(\begin{array}{c|c} b_{11}A & b_{12}A \\ \hline b_{21}A & b_{22}A \end{array}\right) \tag{141}$$

it is obvious that the character of the tensor-product of two representations is the product of the characters which can be written as follows:

$$\chi^{D^{j_1}\otimes_{\mathrm{i}} D^{j_2}}(U_\lambda) \;=\; \chi^{D^{j_1}}(U_\lambda) \;\cdot\; \chi^{D^{j_2}}(U_\lambda) \tag{142}$$

$$= \sum_{m_1=-j_1}^{+j_1} \sum_{m_2=-j_2}^{+j_2} \lambda^{2(m_1+m_2)} \tag{143}$$

$$= \sum_{j=|j_1-j_2|}^{|j_1+j_2|} \underbrace{\sum_{m=-j}^{+j} \lambda^{2m}}_{\chi_{D_j}} \tag{144}$$

The last equal-sign requires a bit of thinking and shall be illustrated here for the simple example: $j_1 = 3$ and $j_2 = 2$. Then the terms of the sum in eq. 143 can be summarized in the following matrix and the summation is carried out row after row:

$$\begin{pmatrix}
\lambda^{2[(-3)+(-2)]} & \lambda^{2[(-3)+(-1)]} & \lambda^{2[(-3)+0]} & \lambda^{2[(-3)+1]} & \lambda^{2[(-3)+2]} \\
\lambda^{2[(-2)+(-2)]} & \lambda^{2[(-2)+(-1)]} & \lambda^{2[(-2)+0]} & \lambda^{2[(-2)+1]} & \lambda^{2[(-2)+2]} \\
\lambda^{2[(-1)+(-2)]} & \lambda^{2[(-1)+(-1)]} & \lambda^{2[(-1)+0]} & \lambda^{2[(-1)+1]} & \lambda^{2[(-1)+2]} \\
\lambda^{2[0+(-2)]} & \lambda^{2[0+(-1)]} & \lambda^{2[0+0]} & \lambda^{2[0+1]} & \lambda^{2[0+2]} \\
\lambda^{2[1+(-2)]} & \lambda^{2[1+(-1)]} & \lambda^{2[1+0]} & \lambda^{2[1+1]} & \lambda^{2[1+2]} \\
\lambda^{2[2+(-2)]} & \lambda^{2[2+(-1)]} & \lambda^{2[2+0]} & \lambda^{2[2+1]} & \lambda^{2[2+2]} \\
\lambda^{2[3+(-2)]} & \lambda^{2[3+(-1)]} & \lambda^{2[3+0]} & \lambda^{2[3+1]} & \lambda^{2[3+2]}
\end{pmatrix} \tag{145}$$

Going from eq. 143 to eq. 144 simply changes the order in which the terms of the previous matrix are added as illustrated in the following figure:

$$\begin{pmatrix} \lambda^{2[(-3)+(-2)]} & \lambda^{2[(-3)+(-1)]} & \lambda^{2[(-3)+0]} & \lambda^{2[(-3)+1]} & \lambda^{2[(-3)+2]} \\ \lambda^{2[(-2)+(-2)]} & \lambda^{2[(-2)+(-1)]} & \lambda^{2[(-2)+0]} & \lambda^{2[(-2)+1]} & \lambda^{2[(-2)+2]} \\ \lambda^{2[(-1)+(-2)]} & \lambda^{2[(-1)+(-1)]} & \lambda^{2[(-1)+0]} & \lambda^{2[(-1)+1]} & \lambda^{2[(-1)+2]} \\ \lambda^{2[0+(-2)]} & \lambda^{2[0+(-1)]} & \lambda^{2[0+0]} & \lambda^{2[0+1]} & \lambda^{2[0+2]} \\ \lambda^{2[1+(-2)]} & \lambda^{2[1+(-1)]} & \lambda^{2[1+0]} & \lambda^{2[1+1]} & \lambda^{2[1+2]} \\ \lambda^{2[2+(-2)]} & \lambda^{2[2+(-1)]} & \lambda^{2[2+0]} & \lambda^{2[2+1]} & \lambda^{2[2+2]} \\ \lambda^{2[3+(-2)]} & \lambda^{2[3+(-1)]} & \lambda^{2[3+0]} & \lambda^{2[3+1]} & \lambda^{2[3+2]} \end{pmatrix} \tag{146}$$

4.3 Clebsch-Gordan coefficients

Every irreducible representation D^l on $\mathcal{V}$ has a canonical basis $\mathcal{B}_{\mathcal{V}} = \{e_l^m\}_{m=-l}^{+l}$ and similarly: every irreducible representation D^s on $\mathcal{W}$ has a canonical basis $\mathcal{B}_{\mathcal{W}} = \{e_s^m\}_{m=-s}^{+s}$. If one reduces the representation $D^l \otimes_{\mathrm{i}} D^s$ on $\mathcal{V} \otimes \mathcal{W}$ into $D^l \otimes_{\mathrm{i}} D^s = \oplus_{j=|l-s|}^{|l+s|} D^j$ then a natural question is how to find a new basis $\mathcal{B}_{\mathcal{V}\otimes\mathcal{W}} = \{e_j^m\}_{m=-j}^{+j}$ on which D^j acts canonically (by canonical we mean that $J_3 e_j^m = m e_j^m$ and $J_+ e_j^j = 0$). This question is answered by the Clebsch-Gordan Coefficients which are the matrix elements of the basis transformation:

$$e_j^m = \sum_{m_l=-l}^{+l} \sum_{m_s=-s}^{+s} \underbrace{\langle e_j^m, e_l^{m_l} \otimes e_s^{m_s} \rangle}_{(lm_l\, sm_s|jm)} e_l^{m_l} \otimes e_s^{m_s} \tag{147}$$

The coefficients $(lm_l\, sm_s|jm)$ in this equation are called the **Clebsch-Gordan coefficients** and can be looked up in tabular form in Quantum Mechanics books.