Selected Solutions for *An Introduction to Tensors and Group Theory for Physicists, 2nd ed.*

March 23, 2019

Introduction

This is an incomplete, evolving solutions manual to [Jee15]. Note that this is the second edition of this text, and that exercise and problem numbering differs between editions. Solutions posted here have primarily been written by students in courses I taught in the UC Berkeley Physics department between 2012 and 2016. If you are interested in improving existing solutions or adding new ones of your own, please email me at nadirj@princeton.edu.

1 A Quick Introduction to Tensors

2 Vector Spaces

Exercise 2.1 : Verify equation 2.7, which states

$$\Delta_{S^2} Y(\theta, \phi) = -l(l-1)Y(\theta, \phi).$$

Solution: Pick a vector $f \in \mathcal{H}_l(\mathbb{R}^3)$. This function has two useful properties: in spherical coordinates, we may write our function as $f(r, \theta, \phi) = r^l Y(\theta, \phi)$ – this is a property of $P_l(\mathbb{R}^3)$, which contains $\mathcal{H}_l(\mathbb{R}^3)$ as a subspace. And, since we have chosen it from $\mathcal{H}_l(\mathbb{R}^3)$, it satisfies $\Delta f = 0$. We write the Laplacian explicitly and combine terms

$$\begin{aligned}\Delta f &= \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Delta_{S^2}\right) r^l Y(\theta, \phi) \\ &= l(l-1)r^{l-2}Y + \frac{2}{r} l r^{l-1} Y + r^{l-2}\Delta_{S^2} Y \\ &= l(l-1)r^{l-2}Y + 2l r^{l-2} Y + r^{l-2}\Delta_{S^2} Y \\ &= l(l+1)r^{l-2}Y + r^{l-2}\Delta_{S^2} Y\end{aligned}$$

But we know this quantity is equal to zero. Therefore,

$$\Delta_{S^2} Y(\theta, \phi) = -l(l+1)Y(\theta, \phi)$$

as claimed. Thus all harmonic l-degree polynomials correspond to a matching spherical harmonic, which corresponds to the restriction of the said polynomial to the unit sphere.

Exercise 2.3 : Given a vector v and a finite basis $\mathcal{B} = \{e_i\}_{i=1,\ldots,n}$, show that the expression of v as a linear combination of the e_i is unique.

Solution: Suppose that our representation of v in the basis $\mathcal{B}$ is explicitly given by $v = a^i e_i$, where the a^i are scalars in the appropriate field. Now, suppose that there exists another representation of v in the same basis, given by $v = b^i e_i$. Then we may write:

$$\begin{aligned}v - v &= 0 \\ &= a^i e_i - b^i e_i \\ &= (a^i - b^i)e_i\end{aligned}$$

From the definition of a basis, our vectors $e_1, e_2, \ldots e_n$ must be linearly independent. Any representation of the zero vector as a linear combination of linearly independent vectors must be the trivial representation, in which all scalar coefficients are 0. Then $(a^i - b^i) = 0$ and therefore $a^i = b^i$. Then $v = b^i e_i$ is not distinct, and $a^i = b^i$ uniquely represents v in the basis $\mathcal{B}$.

Exercise 2.4 : Let $S_n(\mathbb{R}), A_n(\mathbb{R})$ be the sets of n x n symmetric and antisymmetric matrices, respectively. Show that both are real vector spaces, compute their dimensions, and check that $\dim S_n(\mathbb{R}) + \dim A_n(\mathbb{R}) = \dim M_n(\mathbb{R})$, as expected.

Solution: It is relatively simple to show that both $S_n(\mathbb{R})$ and $A_n(\mathbb{R})$ are vector spaces as they are subsets of $M_n(\mathbb{R})$. In fact, we only need to demonstrate that they contain the zero matrix, and are closed under addition and scalar multiplication as all the other properties of a vector space are inherited from $M_n(\mathbb{R})$.

The zero matrix A with components $A_{ij} = 0 \ \forall i, j$ is both symmetric and antisymmetric since $A_{ij} = A_{ji} = -A_{ji}$.

Closure under addition and scalar multiplication holds as $\forall A, B \in S_n(\mathbb{R})$ since

$$cA + B = c\sum_{ij} A_{ij} + \sum_{ij} B_{ij} = c\sum_{ij} A_{ji} + \sum_{ij} B_{ji}$$

as long as $c \in \mathbb{R}$. This also holds in the same manner if A and B are antisymmetric matricies.

Using the given bases for both $S_n(\mathbb{R})$ and $A_n(\mathbb{R})$ all we need to do is count the number of vectors to determine the dimension of each. Recall: E_{ij} is a matrix with a 1 in the ith row, jth column and zeros elsewhere.

We are given that

$$\mathcal{B} \equiv \bigcup_{i<j}\{E_{ij} + E_{ji}\}\bigcup_{i}\{E_{ii}\}$$

is a basis for $S_n(\mathbb{R})$. This basis consists of $\sum_1^n i = \ n(n+1)/2 = n(n-1)/2 + n$ elements which is therefore equal to the dimension of the set of $n \times n$ symmetric matricies.

We are also given that $\bigcup_{i<j}\{E_{ij} - E_{ji}\}$ is a basis for $A_n(\mathbb{R})$. This basis consists of $\sum_1^{n-1} i = n(n-1)/2$ elements which is therefore equal to the dimension of the set of $n \times n$ symmetric matrices.

The sum of the dimensions of S_n and A_n is therefore $n(n-1) + n = n^2$, as expected.

Exercise 2.6 : Using the matrices S_{ij} and A_{ij} from Example 2.9, construct a basis for $H_n(\mathbb{C})$ and compute its dimension.

Solution: Let $H \in H_n(\mathbb{C})$. Then we can decompose H into a real symmetric piece and imaginary antisymmetric piece as

$$H = \tfrac{1}{2}(H + H^T) - \tfrac{i}{2}(H - H^T).$$

This shows that the set

$$\mathcal{B} \equiv \{S_{jk}\}_{j \le k} \bigcup i\{A_{jk}\}_{j<k}$$

spans $H_n(\mathbb{C})$. You can either check directly or take it as given from the previous problem that these vectors are linearly independent, thus proving that $\mathcal{B}$ is a basis. The dimension is then

$$\dim H_n(\mathbb{C}) = \dim S_n(\mathbb{R}) + A_n(\mathbb{R}) = n^2.$$

Exercise 2.8 : Suppose V is finite dimensional and let $T \in \mathcal{L}(V)$. Show that T being one-to-one is equivalent to T being onto.

Solution: Let $T \in \mathcal{L}(V)$ and let $\{e_1, ..., e_n\}$ be a basis for V. We will prov e that

$$T \text{ is 1-1} \Leftrightarrow \{Te_1, ..., Te_n\} \text{ is a basis for } V \Leftrightarrow T \text{ is onto.}$$

Let us first prove the following Lemma.
Lemma 1: $T(0) = 0$.

Proof:

$$\begin{aligned} 0+0=0 &\Rightarrow T(0+0)=T(0) \\ &\Rightarrow T(0)+T(0)=T(0) \text{ by additivity of linear operators} \\ &\Rightarrow T(0)=0 \text{ subtracting } T(0) \text{ on both sides.} \end{aligned}$$

Let us now prove the equivalences.
Step 1: T is 1-1 $\Leftrightarrow \{Te_1, ..., Te_n\}$ is a basis for V.

Assume T is 1-1. Now, suppose $c_1Te_1 + ...c_nTe_n = 0$ for some scalars $c_1, .., c_n$. Then,

$$\begin{aligned} T(c_1e_1+...+c_ne_n)=0=T(0) \text{ (by Lemma 1)} &\Rightarrow c_1e_1+...+c_ne_n=0 \text{ (as } T \text{ is 1-1)} \\ &\Rightarrow c_1=...=c_n=0 \text{ (as } \{e_1,...,e_n\} \text{ is Linearly Independent)} \\ &\Rightarrow \{Te_1,...,Te_n\} \text{ is Linearly Independent of size } n \\ &\Rightarrow \{Te_1,...Te_n\} \text{ is a basis for } V \text{ of dimension } n. \end{aligned}$$

This proves the forward implication. To prove the reverse, assume that $\{Te_1, ..., Te_n\}$ is a basis for V. Then, the set is in particular, linearly independent. Now, suppose $Tu = Tv$ for $u, v \in V$. Then, expanding u, v in terms of the basis $\{e_1, ...e_n\}$, gives

$$\begin{aligned} T(a_1e_1+...+a_ne_n)=T(b_1e_1+...b_ne_n) &\Rightarrow a_1Te_1+...+a_nTe_n=b_1Te_1+...+b_nTe_n \\ &\Rightarrow (a_1-b_1)Te_1+...+(a_n-b_n)Te_n=0 \\ &\Rightarrow a_i-b_i=0 \text{ (as } \{Te_1,...,Te_n\} \text{ is Linearly Independent)} \\ &\Rightarrow u=v \\ &\Rightarrow T \text{ is 1-1.} \end{aligned}$$

Thus, T is 1-1 $\Leftrightarrow \{Te_1, ..., Te_n\}$ is a basis for V.

Step 2: $\{Te_1, ..., Te_n\}$ is a basis for $V \Leftrightarrow T$ is onto.

Assume $\{Te_1, ..., Te_n\}$ is a basis for V. Pick arbitrary $v \in V$. Then, by expanding in terms of the basis $\{Te_1, ..., Te_n\}$, we have,

$$v = c_1Te_1 + ... + c_nTe_n = T(c_1e_1 + ... + c_ne_n) = T(u)$$

for $u = c_1e_1 + ... + c_ne_n \in V$. Thus, for every $v \in V$, there exists a pre-image under T in V. Thus, T is onto. This proves the forward implication.

To prove the reverse, suppose T is onto. Pick arbitrary $v \in V$. Then, there must exist some $u \in V$, such that $T(u) = v$, as T is onto. Now, expanding u in terms of the basis $\{e_1, ..., e_n\}$ gives,

$$v = T(u) = T(c_1e_1 + ... + c_ne_n) = c_1Te_1 + ... + c_nTe_n.$$

Thus, $v \in \text{span}\{Te_1, ..., Te_n\}$. As v was arbitrary, $\{Te_1, ..., Te_n\}$ spans V and since the list has size equal to the dimension of V, $\{Te_1, ..., Te_n\}$ must be a basis for V.

Thus, $\{Te_1, ..., Te_n\}$ is a basis for $V \Leftrightarrow T$ is onto. This proves that for finite dimensional V and any $T \in \mathcal{L}(V)$, T being 1-1 is equivalent to T being onto. This concludes the proof.

Exercise 2.9 : Suppose $T(v) = 0 \Rightarrow v = 0$. Show that this statement is equivalent to T being 1-1, which is, by Exercise 2.8, equivalent to T being onto and hence invertible.

Solution: Assume $T(v) = 0 \Rightarrow v = 0$. Then, for any $x, y \in V$,

$$\begin{aligned} T(x)=T(y) &\Rightarrow T(x)-T(y)=0 \\ &\Rightarrow T(x-y)=0 \text{ (by linearity)} \\ &\Rightarrow x-y=0 \text{ (by the hypothesis)} \\ &\Rightarrow x=y \\ &\Rightarrow T \text{ is 1-1.} \end{aligned}$$

Thus, $T(v) = 0 \Rightarrow v = 0$ implies that T is 1-1.
Now, assume that T is 1-1. Then, for any $v \in V$,

$$T(v) = 0 = T(0) \text{ (by Lemma 1)} \Rightarrow v = 0 \text{ (as } T \text{ is 1-1).}$$

Thus, T being 1-1 implies $T(v) = 0 \Rightarrow v = 0$. Hence, the statement $T(v) = 0 \Rightarrow v = 0$ is equivalent to saying that T is 1-1. This concludes the proof.

Exercise 2.12 : By carefully working with the definitions, show that the e^i defined in (2.25) and satisfying (2.27) are linearly independent.

Solution: The e^i, a set of vectors that span V^*, are defined by

$$e^i(v) = v^i. \tag{1}$$

They are said to be dual to the e_i, i.e.

$$e^i(e_j) = \delta^i_j. \tag{2}$$

To check if they are linearly independent, we must show that

$$(c_i e^i = 0) \Rightarrow c_i = 0 \; \forall \, i. \tag{3}$$

Notice that since e^i are linear functionals, the 0 in equation (3) is the *zero functional*, the function that sends all vectors to the number zero. Since both sides of equation (3) are functions, we can act each side on a vector and see what happens. The most general vector to act our equation on is an arbitrary basis vector e_j where j is a free index, so e_j can represent *any* basis vector $e_i \in \{e_i\}_{i=1,\ldots,n}$, which our e^i are dual to. Acting both sides of equation (3) we find:
$c_i \; e^i \; (e_j)$=$0(e_j)$, which by equation (2) $\Rightarrow c_i \delta^i_j = 0$.
Since δ^i_j will be zero unless $c_i = c_j$ (i.e. $i = j$), we find

$$c_j = 0. \tag{4}$$

This is true for *any* j, since before we picked an arbitrary basis vector e_j with a free index j to act equation (3) on. Therefore, if you give me any j (e.g. 3, 27, or some number over 9,000) I will tell you that it is zero. In math language, you sometimes hear "it is true for any j, and therefore it is true for all j".
Regardless of how you say it, equation (4) shows that equation (3) is true, and so our set of $\{e^i\}_{i=1,\ldots,n}$ are linearly independent. This means that in addition to spanning V^*, they form a basis for it!

Exercise 2.15 : Let A, B $\in M_n(\mathbb{C})$. Define $(\cdot|\cdot)$ on $M_n(\mathbb{C})$ by

$$(A|B) = \frac{1}{2} Tr(A^\dagger B). \tag{5}$$

Check that this is indeed an inner product. Also check that the basis $\{I, \sigma_x, \sigma_y, \sigma_z\}$ for $H_2(\mathbb{C})$ is orthonormal with respect to this inner product.

Solution: An *inner product* is a non-degenerate Hermitian form which, in addition to meeting the basic criteria for being defined as such, obeys

$$(v|v) > 0 \text{ for all } v \in V, v \neq 0 \text{ (positive-definiteness).} \tag{6}$$

First, we must check whether $(A|B) = \frac{1}{2} Tr(A^\dagger B)$ is a non-degenerate Hermitian form. To assist this check, we see what the matrix that we act the trace on in equation (5) looks like: (NOTE: in these matrices, complex conjugates are denoted as C^* rather than $\overline{C}$)
$(A^\dagger|B) =$

$$\begin{pmatrix} A^*_{11} & A^*_{21} & \cdots & A^*_{n1} \\ A^*_{12} & A^*_{22} & \cdots & A^*_{n2} \\ \vdots & & \ddots & \vdots \\ A^*_{1n} & A^*_{2n} & \cdots & A^*_{nn} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2n} \\ \vdots & & \ddots & \vdots \\ B_{n1} & B_{n2} & \cdots & B_{nn} \end{pmatrix}.$$

Multiplying the matrices together, we find $(A^\dagger|B) =$

$$\begin{pmatrix} \sum\limits_{i=1}^{n} A_{i1}^* B_{i1} & \sum\limits_{i=1}^{n} A_{i1}^* B_{i2} & \cdots & \sum\limits_{i=1}^{n} A_{i1}^* B_{in} \\ \sum\limits_{i=1}^{n} A_{i2}^* B_{i1} & \sum\limits_{i=1}^{n} A_{i2}^* B_{i2} & \cdots & \sum\limits_{i=1}^{n} A_{i2}^* B_{in} \\ \vdots & & \ddots & \vdots \\ \sum\limits_{i=1}^{n} A_{in}^* B_{i1} & \sum\limits_{i=1}^{n} A_{in}^* B_{i2} & \cdots & \sum\limits_{i=1}^{n} A_{in}^* B_{in} \end{pmatrix}.$$

Therefore,

$$Tr(A^\dagger B) = \sum_{j=1}^{n}\sum_{i=1}^{n} A_{ij}^* B_{ij}. \tag{7}$$

Now, let's check if equation (5) is a non-degenerate Hermitian form. First, let's check linearity in the second argument, i.e. whether it satisfies Tr($A^\dagger$(B+cD)) = Tr($A^\dagger$B+c$A^\dagger$D)
By equation (7), this is

$\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{n}(A_{ij}^* B_{ij} + cA_{ij}^* D_{ij}) = \sum\limits_{j=1}^{n}\sum\limits_{i=1}^{n} A_{ij}^* B_{ij} + c\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{n} A_{ij}^* D_{ij} =$
Tr($A^\dagger$B)+cTr($A^\dagger$D).

So, yes, equation (5) is linear in the second argument.

NOTE: in the following check for Hermiticity, I will denote complex conjugates as $\overline{C}$ rather than C^* as in the matrices above. Sorry for any confusion.
Second, we must check the Hermiticity of equation (5), i.e. whether it satisfies $\overline{Tr(A^\dagger B)}$=Tr($B^\dagger$A).

Since the trace is just a sum, and $\overline{\sum_i x_i} = \overline{x_1} + \overline{x_2} + \cdots$, we can write $\overline{Tr(A^\dagger B)}$ as Tr($\overline{A^\dagger B}$).

Additionally, $\overline{XY}$=$\overline{X}\ \overline{Y}$, so Tr($\overline{A^\dagger B}$)=Tr($\overline{A^\dagger}\ \overline{B}$). Now, $A^\dagger$ is $\overline{A^T}$, so $\overline{A^\dagger}$= $\overline{\overline{A^T}}$=$A^T$, and $B^\dagger$=$\overline{B^T}$.
Finally, we have:
Tr($A^T\ \overline{B}$) = Tr($\overline{B^T}$A).
Using one last matrix property, $(XY)^T$=$Y^T X^T$, and noting $(A^T\overline{B})^T$=$\overline{B^T}$A, we have:
Tr($(\overline{B^T}A)^T$)=Tr($\overline{B^T}$A).

Since the trace only sums over the diagonal elements, Tr($(XY)^T$)=Tr(XY). Therefore, we have shown that equation (5) satisfies Hermiticity.
Lastly, we must check non-degeneracy of our Hermitian form, i.e. whether for each $v \neq 0 \in V$, there exists $w \in V$ such that $(v|w) \neq 0$. As discussed in the text, this follows as a consequence of positive-definiteness (6), so we check that instead.

From equation (7) we can see that $(A|B)$ takes the form $\frac{1}{2}\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{n} A_{ij}^* B_{ij}$. Since for an inner product we are proving equation (6), we set $B = A$ and find

$$(A|A) = \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n} A_{ij}^* A_{ij} = \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n} |A_{ij}|^2, \tag{8}$$

which is always >0. Therefore, $(A|B)$ is an inner product.
The problem asks us to confirm that $\{I, \sigma_x, \sigma_y, \sigma_z\}$ for $H_2(\mathbb{C})$ is orthonormal with respect to this inner product. For this to be true, it must hold that

$$(e_i|e_j) = \pm\delta_{ij}. \tag{9}$$

Our matrices are

$$\text{I}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Note $\sigma_z^\dagger$=σ_z, $\sigma_y^\dagger$=σ_y, $\sigma_x^\dagger$=σ_x, and $I^\dagger$=I, so if we verify equation (9) for $(A|B)$, then we have satisfied it for $(B|A)$.
Let's get started:

$$(I|I)=\frac{1}{2}\text{Tr}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}=1$$

$$(I|\sigma_x)=\frac{1}{2}\text{Tr}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}=0$$

$$(I|\sigma_y)=\frac{1}{2}\text{Tr}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}=0$$

$$(I|\sigma_z)=\frac{1}{2}\text{Tr}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}=0$$

$$(\sigma_x|\sigma_x)=\frac{1}{2}\text{Tr}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}=1$$

$$(\sigma_x|\sigma_y)=\frac{1}{2}\text{Tr}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}=0$$

$$(\sigma_x|\sigma_z)=\frac{1}{2}\text{Tr}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}=0$$

$$(\sigma_y|\sigma_y)=\frac{1}{2}\text{Tr}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}=1$$

$$(\sigma_y|\sigma_z)=\frac{1}{2}\text{Tr}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}=0$$

$$(\sigma_z|\sigma_z)=\frac{1}{2}\text{Tr}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}=1$$

So, yes, $\{I, \sigma_x, \sigma y, \sigma z\}$ for $H_2(\mathbb{C})$ is orthonormal with respect to equation (5).

Exercise 2.17 : This exercise asks us to show that the definition for a Hilbert space basis is equivalent to our earlier definition for an orthonormal basis when considering a finite-dimensional inner product space V.

Solution: An orthonormal set $\{e_i\}$ is considered an orthonormal Hilbert space basis if:

$$(e_i, f) = 0\ \forall i \quad \Rightarrow \quad f = 0 \tag{10}$$

Our earlier definition for a basis of a vector space V was a linearly independent set $\{e_i\}$ which satisfies Span$\{e_i\}$ = V.

We will first show that the condition for being a Hilbert space basis implies our earlier definition of a basis. Take an orthonormal set $\{e_i\}$ in a finite-dimensional vector space V that has property (10). First we show that the orthogonality of the set implies that it is linearly independent.

If we have a linear combination of the vectors in our orthogonal set such that $c^i e_i = 0$, we can take the inner product of both sides of this equation with another arbitrary basis vector e_j:

$$\begin{aligned} (e_j, c^i e_i) &= (e_j, 0) \\ c^i(e_j, e_i) &= 0 \\ c^i \delta_{ij} &= 0 \quad \text{by orthogonality} \\ c^j = 0 & \end{aligned}$$

Because j was an arbitrary choice, this shows that if $c^i e_i = 0$, then all c^i must be zero. This is precisely the condition for linear independence of the set $\{e_i\}$, so orthogonality implies linear independence.

Now to show that this set spans V, we will consider if there exists any nonzero vector $v \epsilon$ V that is not in the span of $\{e_i\}$.

If there exists a vector v that is not in the span of $\{e_i\}$ then we have

$$v' \equiv v - \sum_j (e_j, v) e_j \neq 0$$

However if we consider taking (e_i, v') for all i, we find that:

$$\begin{aligned}
(e_i, v') &= (e_i, v - \sum_j (e_j, v) e_j) \\
&= (e_i, v) - (e_i, \sum_j (e_j, v) e_j) \\
&= (e_i, v) - \sum_j (e_j, v)(e_i, e_j) \\
&= (e_i, v) - \sum_j (e_j, v) \delta_{ij} \\
&= (e_i, v) - (e_i, v) = 0
\end{aligned}$$

Thus we have shown that $(e_i, v') = 0 \, \forall i$, which by (1) implies that $v' = 0$. This contradicts (2), which implies that our assumption that $v \notin \text{Span}\{e_i\}$ must be false. Thus any vector $v \in V$ is in $\text{Span}\{e_i\}$, therefore $\{e_i\}$ is a basis.

We will now show that having an orthonormal set $\{e_i\}$ such that $\text{Span}\{e_i\} = V$ implies condition (1). First we assume that $(e_i, f) = 0 \, \forall i$ for some $f \in V$. Using the orthogonality of the set, for each i we can write:

$$(e_i, f) = (e_i, f^j e_j) = f^j (e_i, e_j) = f^j \delta_{ij} = f^i = 0$$

Thus $f = 0$. So any orthonormal spanning set also satisfies the criteria for a Hilbert space basis in a finite dimensional inner product space.

Exercise 2.19 : Verify that $P(\mathbb{R})$ is a (real) vector space. Then show that $P(\mathbb{R})$ is infinite-dimensional by showing that, for any finite set $S \subset P(\mathbb{R})$, there is a polynomial that is not in Span S.

Solution: To verify that $P(\mathbb{R})$ is a real vector space, we verify the axioms defining vector space on the elements of $P(\mathbb{R})$, which have the form:

$$f(x) = c_0 + c_1 x + c_2 x^2 + ... + c_n x^n = c_i x^i \tag{11}$$

If we choose the element that has all $c_i = 0$ as the zero element and define $-f(x) = (-c_i) x^i$ so that $f + (-f) = 0$, then axioms 3 and 4 are easily satisfied. We can show that $P(\mathbb{R})$ is closed under addition by choosing two elements, $f = c_i x^i$ and $g = b_i x^i$.

$$f + g = c_i x^i + b_i x^i = (c_i + b_i) x^i \tag{12}$$

which is again an element of $P(\mathbb{R})$ because $\mathbb{R}$ is closed under addition. The remaining axioms are shown very easily.

To show that $P(\mathbb{R})$ is infinite-dimensional, we show that for any finite set $S \subset P(\mathbb{R})$, we can find a polynomial that is not in the span of S. If we take the polynomial element of S that has the maximal degree, n, we simply have to construct a polynomial of degree n+1. This polynomial will not be in the span of S, so no finite set can span $P(\mathbb{R})$. Thus $P(\mathbb{R})$ must be infinite-dimensional.

Exercise 2.20 : Verify that applying the Gram-Schmidt process to $S = \{1, x, x^2, x^3\}$ with the inner product

$$(f|g) \equiv \int_{-1}^{1} f(x)g(x)\ dx \qquad f, g \in P(\mathbb{R})$$

yields, up to normalization, the first four Legendre polynomials, as claimed above. Do the same for the Hermite polynomials, using $[a, b] = (-\infty, \infty)$ and $W(x) = e^{-x^2}$

Solution: For the Legendre polyomials P_i,

$$\begin{aligned}
P_0(x) &= \frac{S_0}{(S_0|S_0)} = 1 \\
P_1(x) &= S_1 - \frac{(P_0, S_1)}{(P_0, P_0)} P_0 \\
&= x - \int_{-1}^{1} x dx \\
&= x \\
P_2(x) &= S_2 - \frac{(P_0, S_2)}{(P_0, P_0)} P_0 - \frac{(P_1, S_2)}{(P_1, P_1)} P_1 \\
&= x^2 - \frac{\int_{-1}^{1} x^2 dx}{\int_{-1}^{1} 1 dx} 1 - \frac{\int_{-1}^{1} x^3 dx}{\int_{-1}^{1} x^2} x \\
&= x^2 - \frac{1}{3}
\end{aligned}$$

Similarly for $P_3(x)$.

For the Hermite polynomials we use $W(x) = e^{-x^2}$ and $[a, b] = [-\infty, \infty]$. We apply Gram-Schmidt to the same set to generate the Hermite Polynomials.

$$\begin{aligned}
H_0(x) &= 1 \\
H_1(x) &= x - \frac{\int_{-\infty}^{\infty} x e^{-x^2} dx}{\int_{-\infty}^{\infty} e^{-x^2} dx} 1 \\
H_1(x) &= x \\
H_2(x) &= x^2 - \frac{\int_{-\infty}^{\infty} x^2 e^{-x^2} dx}{\int_{-\infty}^{\infty} e^{-x^2} dx} 1 - \frac{\int_{-\infty}^{\infty} x^3 e^{-x^2} dx}{\int_{-\infty}^{\infty} x^2 e^{-x^2} dx} x \\
&= x^2 - \frac{1}{2}
\end{aligned}$$

Problem 2-9 : This problem builds on Example 2.22 and further explores different bases for the vector space $P(\mathbb{R})$, the polynomials in one variable x with real coefficients.

a) Compute the matrix corresponding to the operator $\frac{d}{dx} \in \mathcal{L}(P(\mathbb{R}))$ with respect to the basis $\mathcal{B} = \{1, x, x^2, x^3, \ldots\}$.

b) Consider the inner product

$$(f|g) \equiv \int_0^\infty f(x)g(x)e^{-x}\ dx$$

on $P(\mathbb{R})$. Apply the Gram-Schmidt process to the set $S = \{1, x, x^2, x^3\} \subset \mathcal{B}$ to get (up to normalization) the first four ***Laguerre Polynomials***

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= -x+1 \\ L_2(x) &= \frac{1}{2}(x^2-4x+2) \\ L_3(x) &= \frac{1}{6}(-x^3+9x^2-18x+6). \end{aligned}$$

These polynomials arise as solutions to the radial part of the Schrödinger equation for the Hydrogen atom. In this case x is interpreted as a radial variable, hence the range of integration $(0, \infty)$.

Solution: a) We can compute the matrix corresponding to the operator $\frac{d}{dx}$ by considering the effect of the operator on the elements of the basis $\mathcal{B}$, and writing the result in components with respect to $\mathcal{B}$.

$$\begin{aligned} \tfrac{d}{dx}(1) &= 0 = (0,0,0,...) \\ \tfrac{d}{dx}(x) &= 1 = (1,0,0,...) \\ \tfrac{d}{dx}(x^2) &= 2x = (0,2,0,...) \\ \tfrac{d}{dx}(x^n) &= nx^{n-1} = (0,...,0,n,0,...) \end{aligned}$$

where in the last line, n is in the $(n-1)^{th}$ position. If we write these vectors as the columns of a matrix it forms the representation of the $\frac{d}{dx}$ operator in our basis:

$$\left[\frac{d}{dx}\right]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & \cdots \\ 0 & 0 & 2 & \cdots & 0 & \cdots \\ 0 & 0 & 0 & \ddots & \vdots & \\ \vdots & \vdots & \vdots & & 0 & \\ & & & & n & \\ & & & & 0 & \\ & & & & \vdots & \ddots \end{pmatrix}$$

b) Applying Gram-Schmidt with $W(x) = e^{-x}$ and $[a,b] = [0,\infty]$ yields

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= x - \frac{\int_0^\infty xe^{-x}dx}{\int_{-\infty}^\infty e^{-x}dx}1 \\ L_1(x) &= x-1 \\ L_2(x) &= x^2 - \frac{\int_0^\infty x^2e^{-x}dx}{\int_0^\infty e^{-x}dx}1 - \frac{\int_0^\infty x^2(x-1)e^{-x}dx}{\int_0^\infty (x-1)^2e^{-x}dx}(x-1) \\ &= x^2-4x+2 \end{aligned}$$

3 Tensors

Exercises

Exercise 3.1 : By choosing suitable definitions of addition and scalar multiplication, show that $\mathcal{T}_s^r(V)$ is a vector space.

Solution: Let T and S be two (r,s) tensors defined on some vector space V, and let c be a scalar drawn from the field over which that vector space is constructed (either $\mathbb{R}$ or $\mathbb{C}$). We shall define tensor addition and scalar multiplication of a tensor as follows:

$$(T+S)(v_1,\ldots,v_r,f_1,\ldots,f_s) = T(v_1,\ldots,v_r,f_1,\ldots,f_s) + S(v_1,\ldots,v_r,f_1,\ldots,f_s);$$
$$(cT)(v_1,\ldots,v_r,f_1,\ldots,f_s) = c\cdot T(v_1,\ldots,v_r,f_1,\ldots,f_s). \tag{13}$$

With these definitions, we now verify the axioms of a vector space for the set of all (r,s) tens ors on V, denoted $\mathcal{T}_s^r(V)$.

First we deal with the addition postulates. Commutativity follows as a simple result of the commutativity of scalar addition. Let $\mathbf{v} = (v_1,\ldots,v_r)$ and $\mathbf{f} = (f_1,\ldots,f_s)$; then

$$(S+T)(\mathbf{v},\mathbf{f}) = S(\mathbf{v},\mathbf{f}) + T(\mathbf{v},\mathbf{f}) = T(\mathbf{v},\mathbf{f}) + S(\mathbf{v},\mathbf{f}) = (T+S)(\mathbf{v},\mathbf{f}).$$

A similar statement holds for associativity. The 0-tensor may be defined as that which takes any combination of inputs and spits out the scalar 0. Then we have

$$(T+0)(\mathbf{v},\mathbf{f}) = T(\mathbf{v},\mathbf{f}) + 0(\mathbf{v},\mathbf{f}) = T(\mathbf{v},\mathbf{f}) + 0 = T(\mathbf{v},\mathbf{f}),$$

showing that we have our identity. Lastly, for the additive inverse, we pick the tensor which spits out the additive inverse of whatever scalar T produces given a certain combination of inputs. Thus,

$$[T+(-T)](\mathbf{v},\mathbf{f}) = T(\mathbf{v},\mathbf{f}) + (-T)(\mathbf{v},\mathbf{f}) = 0 = 0(\mathbf{v}\ ,\mathbf{f}),$$

as required for the additive inverse of a vector. Now for the scalar multiplication postulates. Distributivity #1 follows simply from the distribut ivity property of scalars:

$$[c(T+S)](\mathbf{v},\mathbf{f}) = c\cdot(T+S)(\mathbf{v},\mathbf{f}) = c\cdot[T(\mathbf{v},\mathbf{f}) + S(\mathbf{v},\mathbf{f})] = c\cdot T(\mathbf{v},\mathbf{f}) + c\cdot S(\mathbf{v},\mathbf{f}) = cT + cS.$$

Distributivity #2 has a similar demonstration. Letting c_1, c_2 be scalars, we have

$$[(c_1+c_2)T](\mathbf{v},\mathbf{f}) = (c_1+c_2)\cdot[T(\mathbf{v},\mathbf{f})] = c_1\cdot T(\mathbf{v},\mathbf{f}) + c_2\cdot T(\mathbf{v},\mathbf{f}) = c_1T + c_2T.$$

Obviously the scalar 1 serves as our multiplicative identity. Lastly, we verify the associativity of scalar multiplication:

$$(c_1c_2)T(\mathbf{v},\mathbf{f}) = (c_1c_2)\cdot T(\mathbf{v},\mathbf{f}) = c_1\cdot[c_2T(\mathbf{v},\mathbf{f})] = c_1(c_2T).$$

Exercise 3.2 : Verify that the Levi-Civita tensor as defined by (3.6) really is multilinear.

Solution: Recall that given vectors $u, v, w \in \mathbb{R}^3$, we define their Levi-Civita symbol by

$$\epsilon(u,v,w) = (u\times v)\cdot w,$$

where $\times$ and $\cdot$ are the typical cross and dot products on $\mathbb{R}^3$. We need to check linearity in each slot to show that this tensor is indeed multilinear.

We begin with v. The cross product distributes over linear combinations and so

$$\epsilon(u, v_1 + cv_2, w) = [u\times(v_1+cv_2)]\cdot w = [u\times v_1 + u\times(cv_2)]\cdot w = (u\times v_1)\cdot w + c[(u\times v_2)\cdot w].$$
$$\implies \epsilon(u, v_1+cv_2, w) = \epsilon(u,v_1,w) + c\epsilon(u,v_2,w).$$

The same argument establishes linearity in the u argument. Finally, linearity in the third slot follows from the fact that the dot product has both the necessary distributivity over addition and you can also pull scalars out. Thus,

$$\epsilon(u,v,w_1+cw_2) = \epsilon(u,v,w_1) = c\epsilon(u,v,w_2).$$

This establishes the multilinearity of the Levi-Civita symbol.

Exercise 3.4 : Show that for a metric g on V,

$$g_i{}^j = \delta_i{}^j,$$

so the $(1,1)$ tensor associated to g (via $g!$) is just the identity operator. You will need the components g^{ij} of the inverse metric, defined in Problem 2-8.

Solution: Recall that by definition $g_i{}^j$ is obtained from g_{ij} by raising an index, i.e. $g_i{}^j \equiv g^{j\ell}g_{\ell j}$. However the components $g^{j\ell}$ of the inverse metric are the inverse of the g_{ij}, and so

$$g_i{}^j = g^{j\ell}g_{\ell j} = \delta_i{}^j.$$

This says that the $(1,1)$ tensor (or linear operator) associated to the $(2,0)$ metric g is just the identity operator!

Exercise 3.6 : Show that for any invertible matrix A, $(A^{-1})^T = (A^T)^{-1}$.

Solution: tarting with an invertible matrix A, we multiply it by its inverse to get $AA^{-1} = I$. Then, taking the transpose of both sides, this means

$$(AA^{-1})^T = I^T \Longrightarrow (A^{-1})^T A^T = I.$$

For the matrix A^T, if its product with another matrix is the identity matrix, that other matrix must be its inverse. So therefore, $(A^{-1})^T = (A^T)^{-1}$.

Exercise 3.7 : Show that for a complex inner product space V, the matrix A implementing an orthonormal change of basis satisfies $A^{-1} = A^\dagger$.

Solution: Let $e_1, e_2, ..., e_n$ be the original orthonormal basis, and let $e_{1'}, e_{2'}, ..., e_{n'}$ be the orthonormal basis that we want to transform into. Let the bases transform such that:

$$e_{i'} = A^k_{i'} e_k$$

Then, because these bases are orthonormal we can write:

$$\delta_{i',j'} = (e_{i'}|e_{j'}) = \bar{A}^k_{i'} A^l_{j'} (e_l|e_k) = \sum_{k=1}^{n} \bar{A}^k_{i'} A^k_{j'}$$

Which in matrices reads

$$A^{-1^\dagger} A^{-1} = I$$

so we must have

$$A^{-1^\dagger} = A$$

and

$$A^\dagger = A^{-1}.$$

Exercise 3.9 : Let $B = \{x, y, z\}$, $B' = \{\frac{1}{\sqrt{2}}(x+iy), z, \frac{1}{\sqrt{2}}(x-iy)\}$ be bases for $H_1(\mathbb{R}^3)$ and consider the operator L_z for which matrix expressions were found with respect to both bases in Example 2.15. Fine the numbers $A^{i'}_j$ and $A^j_{i'}$ and use these, along with (3.33) to obtain, $[L_z]_{B'}$ from $[L_z]_B$.

Solution: First let us write down the relations between the two sets of bases, renaming the vectors so that $B = \{x, y, z\} = \{e_1, e_2, e_3\}$ and $B' = \{\frac{1}{\sqrt{2}}(x+iy), z, \frac{1}{\sqrt{2}}(x-iy)\} = \{e_{1'}, e_{2'}, e_{3'}\}$. We then have

$$\begin{aligned} e_{1'} &= \frac{1}{\sqrt{2}}(e_1 + ie_2) \\ e_{2'} &= e_3 \\ e_{3'} &= \frac{1}{\sqrt{2}}(e_1 - ie_2). \end{aligned}$$

Looking at these equations it is easy to see that we must have

$$A^{-1} = \begin{pmatrix} A^1_{1'} & A^1_{2'} & A^1_{3'} \\ A^2_{1'} & A^2_{2'} & A^2_{3'} \\ A^3_{1'} & A^3_{2'} & A^3_{3'} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & 0 & \frac{-i}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} \tag{14}$$

Knowing A^{-1} we can now get A by simply taking the hermitian conjugate, as we saw in the previous exercise.

$$A = \begin{pmatrix} A^{1'}_1 & A^{1'}_2 & A^{1'}_3 \\ A^{2'}_1 & A^{2'}_2 & A^{2'}_3 \\ A^{3'}_1 & A^{3'}_2 & A^{3'}_3 \end{pmatrix} . = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \end{pmatrix} \tag{15}$$

Then, regarding L_z as a (1,1) tensor, and using the tranformation laws derived for tensors we have

$$\begin{aligned} [L_z]_{\beta'} &= A[L_z]_\beta A^{-1} \\ &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & 0 & \frac{-i}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ \frac{i}{\sqrt{2}} & 0 & \frac{i}{\sqrt{2}} \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \end{aligned}$$

Exercise 3.14 : Compute the dimension of $\Im^r_s$

Solution: Let $\Im^r_s$ be the set of all multilinear functions on

$$\underbrace{V \times V \times ... \times V}_{r \text{ times}} \times \underbrace{V^* \times V^* ... \times V^*}_{s \text{ times}}$$

where V is an n dimensional vector space, and V* is its dual. Recall that $\Im^r_s$ is the same as tensor products of the form

$$\underbrace{V^* \otimes V^* \otimes ... \otimes V^*}_{r \text{ times}} \otimes \underbrace{V \otimes V ... \otimes V}_{s \text{ times}}$$

Furthermore, recall that

$$\dim(V \otimes W) = \dim(V) \dim(W)$$

for any vector spaces V and W. Generalizing this result, we can conclude that

$$\dim(\Im^r_s) = \dim(\underbrace{V^* \otimes V^* \otimes ... \otimes V^*}_{r \text{ times}} \otimes \underbrace{V \otimes V ... \otimes V}_{s \text{ times}}) = n^{r+s}.$$

Exercise 3.15 : Let T_1 and T_2 be tensors of type (r_1, s_1) and (r_2, s_2), respectively, on a vector space V. Show that $T_1 \otimes T_2$ can be viewed as an $(r_1 + r_2, s_1 + s_2)$ tensor, so that the tensor product of two tensors is again a tensor, justifying the nomenclature.

Solution: Let the components of T_1 be $T_{i_1,...,i_{r_1}}{}^{j_1,...,j_{s_1}}$, and the components of T_2 be $T_{i_{r_1+1},...,i_{r_1+r_2}}{}^{j_{s_1+1},...,j_{s_1+s_2}}$, and define

$$Z_{i_1,...,i_{r_1}}{}^{j_1,...,j_{s_1}} = T_{i_1,...,i_{r_1}}{}^{j_1,...,j_{s_1}} T_{i_{r_1+1},...,i_{r_1+r_2}}{}^{j_{s_1+1},...,j_{s_1+s_2}}$$

Then

$$T_1 \otimes T_2 = T_{i_1,\ldots,i_{r_1}}{}^{j_1,\ldots,j_{s_1}} T_{i_{r_1+1},\ldots,i_{r_1+r_2}}{}^{j_{s_1+1},\ldots,j_{s_1+s_2}} (e^{i_1} \otimes \ldots \otimes e^{i_{r_1}} \otimes e_{j_1} \otimes \ldots \otimes e_{j_{s_1}} \otimes e^{i_{r_1+1}} \otimes$$

$$\ldots \otimes e^{i_{r_1+r_2}} \otimes e_{j_{s_1+1}} \otimes \ldots \otimes e_{j_{s_1+s_2}})$$

$= Z_{i_1,\ldots,i_{r_1}}{}^{j_1,\ldots,j_{s_1}} e^{i_1} \otimes \ldots \otimes e^{i_{r_1+r_2}} \otimes e_{j_1} \otimes \ldots \otimes e_{j_{s_1+s_2}}$.
But this is, of course, a type $(r_1 + r_2, s_1 + s_2)$ tensor. Thus, we can conclude that the tensor product of two tensors is again a tensor.

Exercise 3.16 : Show that if $\{e_i\}_{i=1\ldots n}$ and $\{e_{i'}\}_{i=1\ldots n}$ are two arbitrary bases that

$$T(v_1, \ldots, v_{r-1}, e_i, f_1, \ldots, f_{s-1}, e^i) = T(v_1, \ldots, v_{r-1}, e_{i'}, f_1, \ldots, f_{s-1}, e^{i'})$$

so that contraction is well-defined.

Solution: Given that $\{e_i\}$ $i = 1, \ldots, n$ and $\{e_{i'}\}$ $i = 1, \ldots, n$ are two arbitrary basis, we need to show that

$$T\left(v_1, \ldots,\ v_{r-1}, e_i, f_1, \ldots, f_{s-1}, e^i\right) = T(v_1, \ldots,\ v_{r-1}, e_{i'}, f_1, \ldots, f_{s-1}, e^{i'}).$$

This shows that contraction is well-defined since the equality shows that contractions are not dependent on the choices of basis.

From equation (3.16) we have $e_i = A_i^{j'} e_{j'}$. Note that the $A_i^{j'}$ are just numbers. Similarly from equation 3.20 we have for the dual basis transformation $e^i = A^i_{j'} e^{j'}$ Note that $A^i_{j'}$ are just numbers as well. We then have

$$T\left(v_1, \ldots,\ v_{r-1}, e_i, f_1, \ldots, f_{s-1}, e^i\right) = T\left(v_1, \ldots,\ v_{r-1}, A_i^{k'} e_{k'}, f_1, \ldots, f_{s-1}, A^i_{j'} e^{j'}\right)$$

$$T\left(v_1, \ldots,\ v_{r-1}, e_i, f_1, \ldots, f_{s-1}, e^i\right) = A_i^{k'} A^i_{j'} T\left(v_1, \ldots,\ v_{r-1}, e_{k'}, f_1, \ldots, f_{s-1}, e^{j'}\right)$$

Equation 3.18 gives that

$$A_i^{k'} A^i_{j'} = \delta^{k'}_{j'}$$

$$T\left(v_1, \ldots,\ v_{r-1}, e_i, f_1, \ldots, f_{s-1}, e^i\right) = \delta^{k'}_{j'} T\left(v_1, \ldots,\ v_{r-1}, e_{k'}, f_1, \ldots, f_{s-1}, e^{j'}\right).$$

Summing up in respect to k' gives

$$T\left(v_1, \ldots,\ v_{r-1}, e_i, f_1, \ldots, f_{s-1}, e^i\right) = T(v_1, \ldots,\ v_{r-1}, e_{k'}, f_1, \ldots, f_{s-1}, e^{k'}),$$

showing that indeed contraction is well-defined.

Exercise 3.17 : Interpret $e^i \otimes e_j$ as a linear operator, and convince yourself that its matrix representation is

$$[e^i \otimes e_j] = E_{ji}.$$

Recall that E_{ji} is one of the elementary basis matrices introduced way back in Example **??**, and has a 1 in the jth row and ith column and zeros everywhere else.

Solution: To compute $[e^i \otimes e_j]$ we first find the components of $e^i \otimes e_j$:

$$(e^i \otimes e_j)_k{}^l = (e^i \otimes e_j)(e_k, e^l) = e^i(e_k) e_j(e^l) = \delta^i_k \delta^l_j$$

If we call this tensor T then $T_i{}^j = 1$ is the only non-vanishing component. From equation (2.15) we see that when we put these components into a matrix, the entry in the ith column and jth row will be 1, with all others zero, but that is just the matrix E_{ji}. Thus, $[e^i \otimes e_j] = E_{ji}$

Exercise 3.18b : Check that $\{\delta(x - x_0)\}_{x_0 \epsilon \mathbb{R}}$ satisfies (2.28).

Solution: Using Dirac notation we can write the basis $\{\delta(x - x_0)\}_{x_0 \epsilon \mathbb{R}}$ as $\{|x\rangle\}_{x \epsilon \mathbb{R}}$. Then (2.28) states that the set $\{|x\rangle\}_{x \epsilon \mathbb{R}}$ is a Hilbert space basis for $\mathcal{H}$ if it satisfies

$$\langle x \,|\psi\rangle = 0 \;\; \forall x \epsilon \mathbb{R} \implies \psi = 0$$

We can expand $|\psi\rangle$ in the $\{|x\rangle\}$ basis as

$$|\psi\rangle = \int_{-\infty}^{\infty} \psi(x')\,|x'\rangle\,\mathrm{d}x'$$

Now working out the inner product above

$$\begin{aligned}\langle x \,|\psi\rangle &= \int_{-\infty}^{\infty} \psi(x')\langle x\,|x'\rangle\,\mathrm{d}x' \\ &= \int_{-\infty}^{\infty} \psi(x')\delta(x-x')\mathrm{d}x' \\ &= \psi(x)\end{aligned}$$

where we have defined $\langle x\,|x'\rangle \equiv \delta(x-x')$. Thus, $\langle x\,|\psi\rangle = \psi(x)$ so we have shown

$$\langle x \,|\psi\rangle = 0 \;\; \forall x \epsilon \mathbb{R} \implies \psi = 0$$

$\implies \{\delta(x-x_0)\}_{x_0 \epsilon \mathbb{R}}$ satisfies (2.28).

Recall that $\delta(x) \notin L^2(\mathbb{R})$. So how is it that $\{\delta(x-x_0)\}_{x_0\epsilon\mathbb{R}}$ can form a basis for $\mathcal{H}$ if the basis vectors themselves are not in $\mathcal{H}$? The short answer is that they don't, but their corresponding dual vectors do. Recall that an orthonormal set $\{e_i\}$ is defined to be a Hilbert space basis if, for all $f \in \mathcal{H}$,

$$(e_i|f) = 0 \;\; \forall i \implies f = 0.$$

Note, however, that this definition uses the $L(e_i) = (e_i|\cdot) \in \mathcal{H}^*$ rather than the e_i directly. Thus, one could interpret this definition as being a condition on a set of *dual* vectors, and this condition is satisfied by the dual vectors $\langle x|$ as shown above. Note that these dual vectors *are* well-defined and are just equal to the Dirac delta functionals $\delta_x \in \mathcal{H}^*$ defined by

$$\begin{aligned}\delta_x : \mathcal{H} &\longrightarrow \mathbb{C} \\ f &\mapsto f(x).\end{aligned}$$

So the upshot is that the $\langle x|$ are well-defined even though the $|x\rangle$ are not, and the former satisfy the definition of a Hilbert space basis when that definition is interpreted to apply to dual vectors.

Exercise 3.25 : What is the polynomial associated to the Euclidean metric tensor $g = \sum_{i=1}^{3} e^i \otimes e^i$? What is the symmetric tensor $S^2(R^2)$ associated to the polynomial x^2y?

Solution: We have

$$g_{ij} = g(e_i, e_j) = \left(\sum_l e^l \otimes e^l\right)(e_i, e_j) = \sum_l e^l(e_i)e^l(e_j) = \sum_l \delta^l_i \delta^l_j = \delta_{ij}.$$

The associated polynomial is thus

$$f = \delta_{ij}x^i x^j = \delta_{11}x^1x^1 + \delta_{22}x^2x^2 + \delta_{33}x^3x^3 = \boxed{x^2 + y^2 + z^2}$$

Similarly, the polynomial x^2y associated to the symmetric tensor Q_{ijk} can be expressed as

$$f = x^2y = Q_{ijk}x^ix^jx^k$$

In this case, Q_{ijk} must be zero except when $Q_{ijk} = Q_{112} = Q_{121} = Q_{211} = 1/3$. You can check that this yields the correct polynomial:

$$Q_{112}x^1x^1x^2 + Q_{121}x^1x^2x^1 + Q_{211}x^2x^1x^1 = \frac{1}{3}(3x^2y) = x^2y$$

Exercise 3.26 : Substitute $f_l(r) = f^l Y(\theta, \phi)$ into the equation

$$\Delta \left(\frac{f_l(r)}{r^{2l+1}} \right) = 0$$

and show that $Y(\theta, \phi)$ must be a spherical harmonic of degree l. Then use (3.64) to show that if f_l is a harmonic polynomial, then the associated symmetric tensor Q_l must be traceless. If you have trouble showing that Q_l is traceless for arbitrary l, try starting with the $l = 2$ (quadrupole) and $l = 3$ (octopole) cases.

Solution:

$$\begin{aligned}
\Delta \left(\frac{f_l(r)}{r^{2l+1}} \right) &= \Delta \left(\frac{r^l Y(\theta, \phi)}{r^{2l+1}} \right) \\
&= \left[\frac{(l+1)(1+2)}{r^{l+3}} - \frac{2(l+1)}{r^{l+3}} \right] Y(\theta, \phi) + \frac{1}{r^{l+3}} \Delta_{S^2} Y(\theta, \phi) = 0
\end{aligned}$$

$$\begin{aligned}
\implies \Delta_{S^2} Y(\theta, \phi) &= -\left[(l+1)(l+2) - 2(l+1) \right] Y(\theta, \phi) \\
&= -l(l+1) Y(\theta, \phi)
\end{aligned}$$

This shows that $(Y(\theta, \phi))$ is a spherical harmonic, and thus that f is a harmonic polynomial (see earlier sections on this correspondence). The implications for the associated symmetric tensor are as follows. For $l = 2$ we have

$$\begin{aligned}
0 = \Delta Q_{ij} x^i x^j &= \sum_k Q_{ij} \frac{\partial}{\partial x^k} \left(\delta^i_k x^j + \delta^j_k x^i \right) \\
&= \sum_k Q_{ij} \left(\delta^i_k \delta^j_k + \delta^j_k \delta^i_k \right) \\
&= \sum_k 2 Q_{kk}
\end{aligned}$$

and so Q_{ij} is traceless. When $l = 3$ we have

$$\begin{aligned}
0 = \Delta Q_{ijk} x^i x^j x^k &= \sum_m Q_{ijk} \frac{\partial}{\partial x^m} \frac{\partial}{\partial x^m} (x^i x^j x^k) \\
&= 3 \sum_m Q_{ijk} \frac{\partial}{\partial x^m} \delta^i_m x^j x^k \quad \text{by symmetry of } Q_{ijk} \\
&= 6 \sum_m Q_{ijk} \delta^i_m \delta^j_m x^k \quad \text{by symmetry of } Q_{ijk} \\
&= 6 \sum_m Q_{mmk} x^k \\
\implies \sum_m Q_{mmk} &= 0 \ \forall \ k \quad \text{since the } x^k \text{ are linearly independent}
\end{aligned}$$

and so the Q_{ijk} are traceless as well.

Exercise 3.27 : Let $T \in \Lambda^r V^*$. Show that if $\{v_1, ..., v_r\}$ is a linearly dependent set then $T(v_1, ..., v_r) = 0$. Use the same logic to show that if $\{f_1, ..., f_r\} \subset V^*$ is linearly dependent, then $f_1 \wedge ... \wedge f_r = 0$. If $\dim V = n$, show that any set of more than n vectors must be linearly dependent, so that $\Lambda^r V = \Lambda^r V^* = 0$ for $r > n$.

Solution: If $\{v_1, ..., v_r\}$ is a linearly dependent set then one of the vectors can be made from a linear combination of the others. As an example let $v_h = a v_g + b v_k$ such that v_g, v_k and v_h are vectors in $\{v_1, ..., v_r\}$ and a and b are constants. Then

$$T(v_1, ..., v_g, v_k, v_h, ...v_r) = T(v_1, ..., v_g, v_k, (a v_g + b v_k), ...v_r).$$

We can then use multi-linearity to get:

$$T(v_1, ..., v_g, v_k, v_h, ...v_r) = aT(v_1, ..., v_g, v_k, v_g, ...v_r) + bT(v_1, ..., v_g, v_k, v_k, ...v_r).$$

Since these are anti-symmetric tensors we know that $T(v_1, ..., v_r) = 0$ if any of the vectors in $\{v_1, ..., v_r\}$ are the same, which is the case for both terms on the RHS of the above equation. Therefore if $\{v_1, ..., v_r\}$ is linearly dependent then $T(v_1, ..., v_r) = 0$.

If $\{f_1, ..., f_r\}$ is linearly dependent, then (for example) there might exist a vector f_h in $\{f_1, ..., f_r\}$ such that $f_h = af_g + bf_k$ where f_g and f_k are also in $\{f_1, ..., f_r\}$ and a and b are constants. Therefore:

$$f_1 \wedge ...f_g \wedge f_k \wedge f_h \wedge ... \wedge f_r = f_1 \wedge ...f_g \wedge f_k \wedge (af_g + bf_k) \wedge ... \wedge f_r.$$

Since the wedge product is distributive we get:

$$f_1 \wedge ...f_g \wedge f_k \wedge f_h \wedge ... \wedge f_r = a(f_1 \wedge ...f_g \wedge f_k \wedge f_g \wedge ... \wedge f_r) + b(f_1 \wedge ...f_g \wedge f_k \wedge f_k \wedge ... \wedge f_r).$$

Since the wedge product of a vector with itself is zero we see that $f_1 \wedge ... \wedge f_r = 0$ if $\{f_1, ..., f_r\}$ is linearly dependent. If $\dim V = n$ and we have a set of vectors which contains more than n vectors than this set of vectors must be linearly dependent since we know that we only need n vectors to span an n dimensional vector space. Since these vectors would be linearly dependent, from the arguments above we see that $\Lambda^r V = \Lambda^r V^* = 0$ for $r > n$.

Exercise 3.30 : Let $\dim V = n$. Show that the dimension of $\Lambda^r V^*$ and $\Lambda^r V$ is $\binom{n}{r} = \frac{n!}{(n-r)!r!}$.

Solution: We know from the book that $\{e^{i_1} \wedge ...e^{ir}\}_{i_1<...<i_r}$ is the basis for $\Lambda^r V$, so we just need to know how many vectors are in $\{e^{i_1} \wedge ...e^{i_r}\}_{i_1<...<i_r}$. Since this is an anti-symmetric tensor we know that each of the e^{i_j} must be distinct, which is enforced by the condition that $i_1 < ... < i_r$. Each basis vector for $\Lambda^r V$ is a $\wedge$ product of r vectors in V. The basis for V will have n vectors so we get to choose r vectors from n, and the number of different ways that we can do this without repeats is given by the binomial coefficient $\binom{n}{r}$. Therefore the dimension of $\Lambda^r V = \Lambda^r V^*$ is $\binom{n}{r} = \frac{n!}{(n-r)!r!}$.

Exercise 3.31 : You may have seen the values of ϵ_{ijk} defined in terms of cyclic and anti-cyclic permutations. The point of this exercise is to make the connection between that definition and ours, and to see to what extent that definition extends to higher dimensions.

a) Check that the ϵ tensor on $\mathbb{R}^3$ satisfies

$$\epsilon_{ijk} = \begin{cases} +1 & \text{if } \{i, j, k\} \text{ is a cyclic permutation of } \{1, 2, 3\} \\ -1 & \text{if } \{i, j, k\} \text{ is an anticyclic permutation of } \{1, 2, 3\} \\ 0 & \text{otherwise.} \end{cases}$$

Thus for three indices the cyclic permutations are the even rearrangements and the anti-cyclic permutations are the odd ones.

b) Consider ϵ on $\mathbb{R}^4$ with components ϵ_{ijkl}. Are all rearrangements of $\{1, 2, 3, 4\}$ necessarily cyclic or anti-cyclic?

c) Is it true that $\epsilon_{ijkl} = 1$ if $\{i, j, k, l\}$ is a cyclic permutation of $\{1, 2, 3, 4\}$?

Solution: a) We need to prove that in three dimensional Euclidean space, the even permutations are equivalent to cyclic permutations and the odd permutations are equivalent to anti-cyclic permutations.

Let's look at all the permutations of $\{1, 2, 3\}$:
1. [cyclic] (1,2,3) can be obtained by 0 transpositions. (even)
2. [cyclic] (2,3,1) can be obtained by 2 transpositions. (even)
3. [cyclic] (3,1,2) can be obtained by 4 transpositions. (even)
4. [anti-cyclic] (3 2 1) can be obtained by 1 transpositions. (odd)
5. [anti-cyclic] (2 1 3) can be obtained by 3 transpositions. (odd)
6. [anti-cyclic] (1 3 2) can be obtained by 5 transpositions. (odd)

These 6 cases include all the permutations of $\{1,2,3\}$. This gives two conclusions. First, even permutations are equivalent to cyclic permutations and odd permutations are equivalent to anti-cyclic permutation. Second, in three dimensional Euclidean space, the permutations are either cyclic or anti-cyclic.

b) No. In four dimensional Euclidean space, it is possible that a permutation is neither cyclic nor anti-cyclic. One example is (1,3,2,4).

c) No, this is not the case for 4 dimensional Euclidean space. For example, (2,3,4,1) is a cyclic permutation but can be obtained by 3 transpositions, and so $\epsilon_{2341} = -1$.

Exercise 3.33 : Derive (3.79). You may need to consult Section 3.2. Also, Verify (3.81a) by performing the necessary matrix multiplication.

Solution: Deriving (3.79) is almost trivial. From the text (3.17) we know the transformation law for the (0,2) tensor α can be written:

$$\alpha^{i'j'} = A^{i'}_k \alpha^{kl} A^{j'}_l .$$

In terms of matrix multiplication, this is just

$$[\alpha]_{\mathcal{B}'} = A[\alpha]_{\mathcal{B}} A^T$$

As for verifying (3.81a), note that this equation just says

$$[\alpha]_{\beta'} = A[\alpha]_{\beta} A^T = \begin{pmatrix} 0 & -\alpha^3 & \alpha^2\cos\theta + \alpha^1\sin\theta \\ \alpha^3 & 0 & \alpha^2\sin\theta - \alpha^1\cos\theta \\ -\alpha^2\cos\theta - \alpha^1\sin\theta & -\alpha^2\sin\theta + \alpha^1\cos\theta & 0 \end{pmatrix}$$

where

$$A = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Verifying this is actually a very simple matrix multiplication exercise. We will begin by writing out the matrices that needs to be multiplied.

$$[\alpha]_{\beta'} = A[\alpha]_{\beta} A^T = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -\alpha^3 & \alpha^2 \\ \alpha^3 & 0 & -\alpha^1 \\ -\alpha^2 & \alpha^1 & 0 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

We start the multiplication from the right, with

$$\begin{pmatrix} 0 & -\alpha^3 & \alpha^2 \\ \alpha^3 & 0 & -\alpha^1 \\ -\alpha^2 & \alpha^1 & 0 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{16}$$

because for matrix multiplications of more than 2 matrices we would have to reduce them a pair at a time.
We take the first column of the second matrix from equation 16 and multiply it into the first matrix.

$$\begin{pmatrix} 0 & -\alpha^3 & \alpha^2 \\ \alpha^3 & 0 & -\alpha^1 \\ -\alpha^2 & \alpha^1 & 0 \end{pmatrix} \begin{pmatrix} \cos\theta \\ -\sin\theta \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha^3\sin\theta \\ \alpha^3\cos\theta \\ -\alpha^2\cos\theta - \alpha^1\sin\theta \end{pmatrix} \tag{17}$$

The same process done with the second and third column are the following:

$$\begin{pmatrix} 0 & -\alpha^3 & \alpha^2 \\ \alpha^3 & 0 & -\alpha^1 \\ -\alpha^2 & \alpha^1 & 0 \end{pmatrix} \begin{pmatrix} \sin\theta \\ \cos\theta \\ 0 \end{pmatrix} = \begin{pmatrix} -\alpha^3\cos\theta \\ \alpha^3\sin\theta \\ -\alpha^2\sin\theta + \alpha^1\cos\theta \end{pmatrix} \tag{18}$$

$$\begin{pmatrix} 0 & -\alpha^3 & \alpha^2 \\ \alpha^3 & 0 & -\alpha^1 \\ -\alpha^2 & \alpha^1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha^2 \\ -\alpha^1 \\ 0 \end{pmatrix} \tag{19}$$

where equation 18 is for the second column and equation 19 is for the third column.

The product between

$$\begin{pmatrix} 0 & -\alpha^3 & \alpha^2 \\ \alpha^3 & 0 & -\alpha^1 \\ -\alpha^2 & \alpha^1 & 0 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is therefore,

$$\begin{pmatrix} \alpha^3 \sin\theta & -\alpha^3 \cos\theta & \alpha^2 \\ \alpha^3 \cos\theta & \alpha^3 \sin\theta & -\alpha^1 \\ -\alpha^2 \cos\theta + \alpha^1 \sin\theta & -\alpha^2 \sin\theta + \alpha^1 \cos\theta & 0 \end{pmatrix}$$

where the resulting matrix is just the results of equations 17, 18, and 19 in order.

Our initial product of:

$$[\alpha]_{\beta'} = A[\alpha]_\beta A^T = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -\alpha^3 & \alpha^2 \\ \alpha^3 & 0 & -\alpha^1 \\ -\alpha^2 & \alpha^1 & 0 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is now reduced to:

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha^3 \sin\theta & -\alpha^3 \cos\theta & \alpha^2 \\ \alpha^3 \cos\theta & \alpha^3 \sin\theta & -\alpha^1 \\ -\alpha^2 \cos\theta + \alpha^1 \sin\theta & -\alpha^2 \sin\theta + \alpha^1 \cos\theta & 0 \end{pmatrix}.$$

We can now perform the same matrix multiplication again and our result would be

$$\begin{pmatrix} 0 & -\alpha^3 & \alpha^2 \cos\theta + \alpha^1 \sin\theta \\ \alpha^3 & 0 & \alpha^2 \sin\theta - \alpha^1 \cos\theta \\ -\alpha^2 \cos\theta - \alpha^1 \sin\theta & -\alpha^2 \sin\theta + \alpha^1 \cos\theta & 0 \end{pmatrix}$$

which is what we are trying to verify.

Exercise 3.35 : Use 3.79 to show the bivector $\tilde{\omega}$ in the body frame is

$$[\tilde{\omega}]_K = A^{-1}\frac{dA}{dt}.$$

Then use this equation and $[\mathbf{v}]_{\mathbf{K}'} = \frac{dA}{dt}A^{-1}[\mathbf{r}]_{\mathbf{K}'}$ to show $\mathbf{v} = \omega \times \mathbf{r}$ in the body frame.

Solution: We can solve for $[\tilde{\omega}]_K$ starting with the matrix transformation law

$$[\tilde{\omega}]_{K'} = A[\tilde{\omega}]_K A^T$$

Right hand multiplying by A, and using the fact that A is an orthogonal matrix, we have

$$[\tilde{\omega}]_{K'} A = A[\tilde{\omega}]_K$$

Left hand multiplying by A^{-1} and rearranging gives

$$[\tilde{\omega}]_K = A^{-1}[\tilde{\omega}]_{K'} A$$

Then using $[\tilde{\omega}]_{K'} = \frac{dA}{dt}A^{-1}$ we get

$$\boxed{[\tilde{\omega}]_K = A^{-1}\frac{dA}{dt}.}$$

Next, to show that $\mathbf{v} = \omega \times \mathbf{r}$ we begin with

$$[\mathbf{v}]_{\mathbf{K}'} = \frac{dA}{dt}A^{-1}[\mathbf{r}]_{\mathbf{K}'}.$$

Left hand multiplying by A^{-1} gives $A^{-1}[\mathbf{v}]_{\mathbf{K}'} = A^{-1}\frac{dA}{dt}A^{-1}[\mathbf{r}]_{\mathbf{K}'}$. Then using $[\mathbf{r}]_{\mathbf{K}'} = A[\mathbf{r}]_{\mathbf{K}}$ and $[\mathbf{v}]_{\mathbf{K}'} = A[\mathbf{v}]_{\mathbf{K}}$ gives

$$[\mathbf{v}]_{\mathbf{K}} = A^{-1}\frac{dA}{dt}[\mathbf{r}]_{\mathbf{K}}.$$

Then we use the definition of the angular velocity vector to be $\omega = J(\tilde{\omega})$ to write the matrix

$$A^{-1}\frac{dA}{dt} = \begin{pmatrix} 0 & -\omega^3 & \omega^2 \\ \omega^3 & 0 & -\omega^1 \\ -\omega^2 & \omega^1 & 0 \end{pmatrix}$$

The we calculate

$$A^{-1}\frac{dA}{dt}[\mathbf{r}]_{\mathbf{K}} = \begin{pmatrix} 0 & -\omega^3 & \omega^2 \\ \omega^3 & 0 & -\omega^1 \\ -\omega^2 & \omega^1 & 0 \end{pmatrix} \begin{pmatrix} r^1 \\ r^2 \\ r^3 \end{pmatrix}$$

$$= \begin{pmatrix} \omega^2 r^3 - \omega^3 r^2 \\ \omega^3 r^1 - \omega^1 r^3 \\ \omega^1 r^2 - \omega^2 r^1 \end{pmatrix} = [\omega \times \mathbf{r}]_{\mathbf{K}}.$$

Thus we have shown $[\mathbf{v}]_{\mathbf{K}} = [\omega \times \mathbf{r}]_{\mathbf{K}}$.

Problems

Problem 3-1 : In this problem we explore the properties of $n \times n$ orthogonal matrices. This is the set of real invertible matrices A satisfying $A^T = A^{-1}$, and is denoted $O(n)$.

a) Is $O(n)$ a vector subspace of $M_n(\mathbb{R})$?

b) Show that the product of two orthogonal matrices is again orthogonal, that the inverse of an orthogonal matrix is again orthogonal, and that the identity matrix is orthogonal. These properties show that $O(n)$ is a ***group***, i.e. a set with an associative multiplication operation and identity element such that the set is closed under multiplication and every element has a multiplicative inverse. Groups are the subject of Chapter 4.

c) Show that the columns of an orthogonal matrix A, viewed as vectors in $\mathbb{R}^n$, are mutually orthogonal under the usual inner product. Show the same for the rows. Show that for an active transformation, i.e.

$$[e_{i'}]_{\mathcal{B}} = A[e_i]_{\mathcal{B}}$$

where $\mathcal{B} = \{e_i\}$ so that

$$[e_i]_{\mathcal{B}}^T = (0, \ldots, \underbrace{1}_{i\text{th slot}}, \ldots, 0),$$

the columns of A *are* the $[e_{i'}]_{\mathcal{B}}$. In other words, the components of the *new* basis vectors in the *old* basis are just the columns of A. This also shows that for a passive transformation, where

$$[e_i]_{\mathcal{B}'} = A[e_i]_{\mathcal{B}}$$

the columns of A are the components of the *old* basis vectors in the *new* basis.

d) Show that the orthogonal matrices A with $|A| = 1$, the rotations, form a subgroup unto themselves, denoted $SO(n)$. Do the matrices with $|A| = -1$ also form a subgroup?

Solution: **(a)** To be a vector subspace, $O(n)$ must contain the zero vector. For the vector space of matrices, the zero vector is simply the zero matrix. But, the zero matrix has no inverse, therefore it cannot satisfy $A^T \neq A^{-1}$, and therefore $O(n)$ is not a vector subspace of $M_n(\mathbb{R})$.
(b) Give two orthogonal matrices A and B, their product is AB. To be orthogonal, we require that $(AB)^{-1} = (AB)^T$. This is easily shown:

$$(AB)^T = B^T A^T = B^{-1}A^{-1} = (AB)^{-1}.$$

To show that the inverse of an orthogonal matrix is orthogonal, for an orthogonal matrix A,

$$A^{-1} = A^T \Longrightarrow (A^{-1})^{-1} = (A^T)^{-1} = (A^{-1})^T$$

So that means $(A^{-1})^{-1} = (A^{-1})^T$, and therefore A^{-1} is an orthogonal matrix. For the identity matrix I, $I^{-1} = I$ and $I^T = I$, so therefore $I^T = I^{-1}$ and so I is an orthogonal matrix.

(c) To show that the column vectors and the row vectors are mutually orthogonal under the usual inner product, we notice that $A^T A = I$. In components, this reads $\sum_{j'} A_i^{j'} A_k^{j'} = \delta_{ik}$. Since $A_i^{j'}$ with i fixed and j' varying is a column of A, this says that the columns of A are orthonormal vectors. Writing down $AA^T = I$ would yield the same conclusion for the rows of A.

For the last part of this problem, we evaluate both sides of the identity $[e_{i'}]_{\mathcal{B}} = A[e_i]_{\mathcal{B}}$ to prove it. For the left hand side, we have

$$\begin{aligned}
[e_{i'}]_{\mathcal{B}} &= \left[\sum_j A_{i'}^j e_j\right]_{\mathcal{B}} \\
&= \begin{pmatrix} A_{i'}^1 \\ A_{i'}^2 \\ \cdot \\ \cdot \\ A_{i'}^n \end{pmatrix}
\end{aligned}$$

Meanwhile,

$$\begin{aligned}
A[e_i]_{\mathcal{B}} &= \begin{pmatrix} A_1^{1'} & A_2^{1'} & \dots & A_n^{1'} \\ A_1^{2'} & A_2^{2'} & \dots & A_n^{2'} \\ \vdots & \vdots & \vdots & \vdots \\ A_1^{n'} & A_{2'}^{n'} & \dots & A_n^{n'} \end{pmatrix} \begin{pmatrix} 0 \\ \cdot \\ 1 \\ \cdot \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} A_{i'}^1 \\ A_{i'}^2 \\ \cdot \\ \cdot \\ A_{i'}^n \end{pmatrix}
\end{aligned}$$

and we're done.

(d) Given two matrices A and B whose determinants are 1, their product is AB. Taking the determinant gives $|AB| = |A||B| = 1 \times 1 = 1$. Therefore, the rotations form a subgroup unto themselves. For the matrices A where $|A| = -1$, given two matrices A and B that satisfy this condition, the determinant of their product is: $|AB| = |A||B| = -1 \times -1 = 1$. Thus, these matrices do *not* form their own subgroup.

Problem 3-2 : In this problem we will compute the dimension of the space of (0,r) symmetric tensors $S^r(V)$.

(a) Let dim $V = n$ and $\{e_i\}_{i=1,\ldots,n}$ be a basis for V. Argue that dim $S^r(V)$ is given by the number of ways you can choose r (possibly repeated) vectors from the basis $\{e_i\}_{i=1,\ldots,n}$.

(b) We have now reduced the problem to combinatorics problem: how many ways can you choose r objects from a set of n objects, where any object can be chosen more than once? The answer is

$$\dim S^r(V) = \binom{n+r-1}{n-1} = \frac{(n+r-1)!}{r!(n-1)!} \tag{20}$$

Solution: (a) Given a symmetric tensor space of rank r over vector space V, $S^r(V)$, with dimension $V = n$, we know than a rank r tensor over this space must be composed of a linear combination of tensor products of r vectors from V. Since we are considering symmetric tensors, we can choose to repeat these vectors.

Consider now a symmetric tensor $T \in S^r(V)$. We can write it as a sum of tensor products using Einstein summation in the following way:

$$T = T^{i_1 \ldots i_r} e_{i_1} \otimes \ldots \otimes e_{i_r}. \tag{21}$$

If T is symmetric then the components $T^{i_1 \ldots i_r}$ are invariant under any permutation of the indices. This means we can factor out the components in the summation in the following way:

$$T = \sum_{i_1 \leq \ldots \leq i_r} T^{i_1 \ldots i_r} \underbrace{\sum_{\text{perm}} e_{i_1} \otimes \ldots \otimes e_{i_r}}_{\text{basis vectors}} \tag{22}$$

where the second sum is over distinct permutations of the r-tuples $(i_1, \cdots, i_r)$. This expression shows that these sums over permutations of tensor products span $S^r(V)$, and it's not hard to check that they in fact form a basis. Each basis vector of $S^r(V)$ is determined by the choice of integers $(i_1, \cdots, i_r)$ subject to the restriction $i_1 \leq \ldots \leq i_r$, or equivalently by a choice of r (possibly repeated) basis vectors from the basis $\{e_i\}_{i=1,\ldots,n}$ of V.
By way of example, we consider a symmetric rank 2 tensor $T \in S^2(\mathbb{R}^3)$:

$$T = T^{ij} e_i \otimes e_j = T^{ij} \frac{1}{2} \left(e_i \otimes e_j + e_j \otimes e_i \right). \tag{23}$$

The basis for the symmetric tensor space is the set $\{e_1 \otimes e^1, e_2 \otimes e_2, e_3 \otimes e_3, e_1 \otimes e_2 + e_2 \otimes e_1, e_1 \otimes e_3 + e_3 \otimes e_1, e_2 \otimes e_3 + e_3 \otimes e_2\}$. Then the associated matrix $[T]$ is simply a 3×3 symmetric matrix with entries from $\mathbb{R}^3$:

$$\begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix} \tag{24}$$

There is a total of 6 free parameters needed to describe the matrix (three for the diagonal terms, these are the repeated terms, and three for the off-diagonal terms, these are the cross terms). This matches the description of the tensors we have above in terms of the tensor products of the basis vectors. It is then clear that the dimension of a symmetric tensor space $S^r(V)$ is the number of ways one can choose r vectors from the basis of V.

(b) We now want to determine how we can count these basis vectors. We will use the 'balls and walls' method to count the dimension. First, we define a set of r balls (this is the rank of our tensor), to be placed into n bins (the dimension of vector space V). Since there is a total of n bins, there must be $n-1$ walls between the bins. In this counting method we are allowed to have empty bins as well as bins that contain more than one ball, because some tensor products will have some of the e_i absent and others repeated. There is a total of $n-1$ walls for n bins. There is also a total of r balls that we can place in any bin. Thus there is a total of $r+n-1$ slots that can be filled (see schematic below). If we keep the balls fixed, and instead choose where to place the $n-1$ walls, we are choosing $n-1$ configurations from a total number of $r+n-1$ slots.

Let us once again consider an example for clarity. Consider $r = 4$ balls and $n = 5$ bins. Note that this corresponds to $S^4(V)$, where $\dim(V) = 5$. First, we line up the balls:

$$\bullet \quad \bullet \quad \bullet \quad \bullet$$

Now, we place walls in the configuration. Note that we can place a wall between any two balls, even if there is another wall there. We might have a result that looks like this:

$$\underline{\bullet} \quad \underline{/} \quad \underline{/} \quad \underline{\bullet} \quad \underline{/} \quad \underline{\bullet} \quad \underline{/} \quad \underline{\bullet}$$

We indicated each slot with a dash below it. To make the connection between our problem consider that each bin corresponds to a basis vector e_i of V. Each ball in a given bin corresponds to that bin's basis vector being part of the tensor product. Thus our schematic above corresponds to the tensor product $e_1 \otimes e_3 \otimes e_4 \otimes e_5$. To make this symmetric we simply permute through all the possible orderings of the basis vectors. Note that all the other orderings will not be counted in this counting method, because the balls are indistinguishable from one another. Our final result is thus the dimension of $S^r(V)$, is given by the number of ways we can arrange $n-1$ walls in $r+n-1$ slots, or

$$\dim S^r(V) = \binom{n+r-1}{n-1} = \frac{(n+r-1)!}{r!(n-1)!} \tag{25}$$

Problem 3-3 : Prove the following basic properties of the determinant directly from the definition (3.72). We will restrict our discussion to operations with columns, though it can be shown that all the corresponding statements for rows are true as well.

a) Any matrix with a column of zeros has $|A| = 0$.

b) Multiplying a column by a scalar c multiplies the whole determinant by c.

c) The determinant changes sign under interchange of any two columns.

d) Adding two columns together, i.e. sending $A_i \to A_i + A_j$ for any i and j, doesn't change the value of the determinant.

Solution: **(a)** The simple way to do this problem uses the multilinearity of tensors, and observing that one can create a column of zeros in a matrix simply by multiplying that column by the scalar, 0. Then, multilinearity lets you pull this scalar out of the argument of the tensor and multiply the entire tensor by this scalar:

$$\epsilon(A_1, ..., cA_i, ..., A_n) = c\epsilon(A_1, ..., A_i, ..., A_n) = c|A| \quad = 0 \quad \text{if} \;\; c = 0 \qquad \square$$

(b) Using the definition of the determinant as $|A| = \epsilon(A_1, ..., A_n)$ and the tensor's multilinearity, if we multiply some column, A_i by a scalar c, we are able to pull out the scalar:

$$\epsilon(A_1, ..., cA_i, ..., A_n) = c\epsilon(A_1, ..., A_i, ..., A_n) = c|A| \quad \square$$

(c) This is readily apparent from the fact that the Levi-Civita tensor is an anti-symmetric tensor, so by definition, an interchange of any set of arguments will reverse the sign in the tensor, thus reversing the sign of the determinant:

$$|A| = \epsilon(A_1, ..., A_i, ..., A_j, ..., A_n)$$

and

$$\epsilon(A_1, ..., A_j, ..., A_i, ..., A_n) = -\epsilon(A_1, ..., A_i, ..., A_j, ..., A_n) = -|A|$$

resulting in a reversal of the sign of the determinant. □

(d) Since,

$$|A| = \epsilon(A_1, ..., A_i, ..., A_j, ..., A_n)$$

we must evaluate

$$\epsilon(A_1, ..., A_i + A_j, ..., A_j, ..., A_n) \tag{26}$$

But, by multilinearity, we know that (26) can be written as

$$\epsilon(A_1, ..., A_i, ..., A_j, ..., A_n) + \epsilon(A_1, ..., A_j, ..., A_j, ..., A_n) = |A| + |A'|$$

but the first property of antisymmetric tensors requires that its value is 0 if any arguments are repeated, so $|A'| = 0$ and the new value of the determinant is the same as the original, $|A|$.

Problem 3-4 : One can extend the definition of determinants from matrices to more general linear operators as follows: We know that a linear operator T on a vector space V (equipped with an inner product and orthonormal basis $\{e_i\}_{i=1...n}$) can be extended to an operator on the p-fold tensor product $\mathcal{T}^0_p(V)$ by

$$T(v_1 \otimes \ldots \otimes v_p) = (Tv_1) \otimes \ldots \otimes (Tv_p)$$

and thus, since $\Lambda^n V \subset \mathcal{T}^0_n(V)$, the action of T extends to $\Lambda^n V$ similarly by

$$T(v_1 \wedge \ldots \wedge v_n) = (Tv_1) \wedge \ldots \wedge (Tv_n).$$

Consider then the action of T on the contravariant version of ϵ, the tensor $\tilde{\epsilon} \equiv e_1 \wedge \ldots \wedge e_n$. We know from Exercise (3.30) that $\Lambda^n V$ is one dimensional, so that $T(\tilde{\epsilon}) = (Te_1) \wedge \ldots \wedge (Te_n)$ is proportional to $\tilde{\epsilon}$. We then define the ***determinant*** of T to be this proportionality constant, so that

$$(Te_1) \wedge \ldots \wedge (Te_n) \equiv |T|\, e_1 \wedge \ldots \wedge e_n. \tag{27}$$

a) Show by expanding the left hand side of (27) in components that this more general definition reduces to the old one of (3.73) in the case of $V = \mathbb{R}^n$.

b) Use this definition of the determinant to show that for two linear operators B and C on V,

$$|BC| = |B||C|.$$

In particular, this result holds when B and C are square matrices.

c) Use b) to show that the determinant of a matrix is invariant under similarity transformations (see Example 3.8). Conclude that we could have defined the determinant of a linear operator T as the determinant of its matrix in *any* basis.

Solution: **(a)**

$$\begin{aligned}
T(\tilde{\epsilon}) = T(e_1) \wedge ... \wedge T(e_n) &= (\sum_{i_1} T_{i_1 1} e_{i_1}) \wedge ... \wedge (\sum_{i_n} T_{i_n n} e_{i_n}) \\
&= (\sum_{i_1,...,1_n} T_{i_1 1} ... T_{i_n n})(e_{i_1} \wedge ... \wedge e_{i_n}) \\
&= (\sum_{i_1,...,1_n} T_{i_1 1} ... T_{i_n n}) \epsilon_{i_1,...,1_n} (e_1 \wedge ... \wedge e_n) \\
&= |\, T\,|\, (e_1 \wedge ... \wedge e_n) \\
&= |\, T\,|\, \tilde{\epsilon}
\end{aligned}$$

(b)

$$\begin{aligned}
|\, BC \,|\, \tilde{\epsilon} &= BC(\tilde{\epsilon}) \\
&= B(C(\tilde{\epsilon})) \\
&= B(C(e_1) \wedge ... \wedge C(e_n)) \\
&= B(|\, C \,|\, \tilde{\epsilon}) \\
&= |\, C \,|\, B(\tilde{\epsilon}) \\
&= |\, C \,||\, B \,|\, \tilde{\epsilon}
\end{aligned}$$

$$\Longrightarrow |\, BC \,| = |\, B \,||\, C \,|$$

(c) From part b we know that the determinant of a product is the product of the determinants:

$$\begin{aligned}
\Longrightarrow |\, I \,| &= |\, AA^{-1} \,| \\
&= |\, A \,||\, A^{-1} \,| \\
|\, A^{-1} \,| &= \frac{1}{|\, A \,|}
\end{aligned}$$

Now we apply the similarity transformation to T:

$$\begin{aligned}
|\, ATA^{-1} \,| &= |\, A \,||\, T \,||\, A^{-1} \,| \\
&= |\, A \,||\, T \,| \frac{1}{|\, A \,|} \\
&= |\, T \,|
\end{aligned}$$

From this we see we could have defined the determinant of T as the determinant of its matrix representation in any basis.

Problem 3-5 : Let V be a vector space with an inner product and orthonormal basis $\{e_i\}_{i=1...n}$. Prove that a linear operator T is invertible if and only if $|T| \neq 0$, as follows:

a) Show that T is invertible if and *only if* $\{T(e_i)\}_{i=1...n}$ is a linearly independent set (see Exercise 2.9 for the 'if' part of the statement).

b) Show that $|T| \neq 0$ if and only if $\{T(e_i)\}_{i=1...n}$ is a linearly independent set.

c) This is not a problem, just a comment. In Example 3.28 we interpreted the determinant of a matrix A as the oriented volume of the n-cube determined by $\{Ae_i\}$. As you just showed, if A is not invertible then the Ae_i are linearly dependent, hence span a space of dimension less than n and thus yield an n-dimensional volume of 0. Thus, the geometrical picture is consistent with the results you just obtained!

Solution: **a)** To show invertibility implies linear independence: From Exercise 2.9, T being invertible implies that T is 1 to 1. So assume T is 1-to-1. Then

$$c^i(Te_i) = 0 \implies T(c^i e_i) = 0 \implies c^i e_i = 0$$

since T is 1 to 1. Because $\{e_i\}_{i=1,...,n}$ is a basis, you know the e_i are LI and so $c^i = 0 \,\forall\, i$ which then means that the $T(e_i)$ are linearly independent as well.

To show linear independence implies invertibility: assume $\{T(e_i)\}_{i=1,...,n}$ is linearly independent. This means $T(c^i e_i) = 0 \implies c^i = 0$. Thus, $T(v) = 0 \implies v = 0$, which is exactly the condition that T is 1 to 1, which $\implies$ invertibility by exercise 2.7

b) For linear independence implies $|T| \neq 0$: assume $\{T(e_i)\}_{i=1,...,n}$ is linearly independent, which by the results of part a means T is invertible so $TT^{-1} = I$. Since $|TT^{-1}| = |T||T^{-1}|$,

$$|TT^{-1}| = 1$$

$$\implies |T||T^{-1}| = 1$$

$$\implies |T^{-1}| = 1/|T|$$

So, $|T| \neq 0$.

For the $|T| \neq 0$ implies linear independence: assume $|T| \neq 0$. If $T(e_i)$ is a linearly dependent set, there is some vector in the set, say e^k, such that $e^k = a_1 e^1 + ... + a_n e^n$. But then we would have

$$\begin{aligned} \epsilon &= e^1 \wedge ... \wedge e^k \wedge ... \wedge e^n \\ &= a_1(e^1 \wedge ... \wedge e^1 \wedge ... \wedge e^n) + ... + a_n(e^1 \wedge ... \wedge e^n \wedge ... \wedge e^n) \end{aligned}$$

and since any vector $\wedge$ itself is zero, $\epsilon = 0$ so $|T| = 0$ if the $T(e_i)$ are linearly dependent. Thus, for $|T| \neq 0$, $\{T(e_i)\}_{i=1,...,n}$ must be linearly independent.

---

4 Groups, Lie Groups, and Lie Algebras

Exercises

Exercise 4.3 : Prove the cancelation laws for groups, ie that:

$$\forall g_1, g_2, h \in G$$
$$(1) \quad g_1 h = g_2 h \implies g_1 = g_2$$
$$(2) \quad h g_1 = h g_2 \implies g_1 = g_2$$

Solution: Multiplying both sides of (1) by h^{-1} on the right yields the following by associativity:

$$g_1(hh^{-1}) = g_2(hh^{-1}) \quad \implies \quad g_1 e = g_2 e$$

Any element in a group multiplied by its identity returns the original element, therefore:

$$g_1 = g_2$$

By multiplying both sides of (2) by h^{-1} on the left, the same result as before can be obtained.

$$(h^{-1}h)g_1 = (h^{-1}h)g_2 \quad \implies \quad eg_1 = eg_2 \quad \implies \quad g_1 = g_2$$

Exercise 4.4 : Show that:

$$[v]^T[w] = [v]^T[T]^T[T][w] \quad \forall v, w \in V$$

If and only if $[T]^T[T] = I$.

Solution: Let $v = e_i, \quad w = e_j$ where $\{e_i\}_{i=1,\ldots,n}$ is orthonormal.
We prove the “only if" part of the statement, since the "if" part is trivial.

Replacing v and w in (4) with the elementary column vector e_i and elementary row vector e_j yields:

$$(4) \quad [e_i]^T[e_j] = [e_i]^T[T]^T[T][e_j]$$

Taking the left side of (4) it can be seen that:

$$[e_i]^T[e_j] = (e_i|e_j) = \delta_{ij}$$

The term on the right being the kronecker delta which has the value 1 when $i = j$ and the value 0 when $i \neq j$.
When an operator T of an isometry set, a set of operators which preserve the dot product between two vectors, is applied to e_i and e_j the following is true:

$$(e_i|e_j) = (Te_i|Te_j) = [Te_i]^T[Te_j] = [e_i]^T[T]^T[T][e_j]$$

It can be seen that the product $[T]^T[T]$ adopts the indices of the elementary vectors which surround it. Due to this the right side of the equation above, which is also the right side of (4), becomes:

$$([T]^T[T])_{ij}$$

This being the case:

$$([T]^T[T])_{ij} = \delta_{ij}$$

$([T]^T[T])_{ij}$, like δ_{ij}, are the indices of the identity matrix, I, therefore:

$$[T]^T[T] = I$$

Exercise 4.5 : Verify that $O(n)$ is a group using the orthogonality condition:

$$T^T = T^{-1} \ .$$

This is the same as Problem 3-1 b.

Solution:

$$\forall T, U, V \in O(n)$$
$$T^T = T^{-1}, \quad U^T = U^{-1}, \quad V^T = V^{-1}$$

The following axioms are proven with the above being true.

(a) Closure by multiplication

$$(TU)^T = U^T T^T = U^{-1}T^{-1} = (TU)^{-1}$$

Therefore TU also satisfies the orthogonality condition and is in $O(n)$.

(b) Associativity
Associativity is a general property that governs the composition of linear operators, including matrix multiplication. Thus, the property carries on to the $O(n)$ set of linear operators.

(c) Existence of the identity
To show that I^T is equal to I^{-1} we can use the fact that the transpose of the identity matrix is equal to itself and the inverse of the identity matrix is equal to itself as well.

$$I^T = I = I^{-1}$$

Therefore:

$$I^T = I^{-1}$$

The identity matrix satisfies the orthogonality condition so it exists in $O(n)$.

(d) Existence of inverses

We want to show that if $T \in O(n)$, then T^{-1} is also in $O(n)$, i.e.

$$(T^{-1})^{-1} = (T^{-1})^T$$

To prove this notice that:

$$(T^{-1})^{-1} = T$$

Now taking the rewritten orthogonality condition:

$$I = T^T T$$

Multiplying both sides by the tranpose of the inverse of T yields:

$$(T^{-1})^T = (T^{-1})^T T^T T = (TT^{-1})^T T = I^T T = IT = T$$

Therefore:

$$(T^{-1})^{-1} = T = (T^{-1})^T \implies (T^{-1})^{-1} = (T^{-1})^T$$

The inverse of T thus satisfies the orthogonality condition and so it exists in $O(n)$.

With all the axioms satisfied by $O(n)$ we have shown that it is indeed a group.

Exercise 4.6 : Verify that U(n) is a group, using the defining condition (4.15).

Solution: $U(n)$ is defined as the set of all $n \times n$ unitary matrices over the complex numbers. Let A be a unitary matrix. By the defining condition, $A^\dagger = A^{-1}$, where $A^\dagger$ is the adjoint.

We will verify that $U(n)$ is a group by going through the four axioms:

1. Closure: Let $A, B \in U(n)$. Then $(AB)^\dagger = B^\dagger A^\dagger$. Using the defining condition, $B^\dagger A^\dagger = B^{-1}A^{-1} = (AB)^{-1} \implies AB \in U(n)$.

2. Associativity: Since $U(n)$ is a subgroup of $GL(n, \mathbb{C})$, it inherits associativity from $GL(n, \mathbb{C})$.

3. Existence of an identity: Let I be the identity matrix. Since I is conjugate-symmetric and its own inverse, $I^\dagger = I^{-1} \implies I \in U(n)$.

4. Existence of inverses: Let $A \in U(n)$. Consider $(A^{-1})^\dagger$. $A^\dagger(A^{-1})^\dagger = (A^{-1}A)^\dagger = I^\dagger = I$ which means $(A^{-1})^\dagger = (A^\dagger)^{-1} = (A^{-1})^{-1} = A$, so therefore $A^{-1} \in U(n)$.

Exercise 4.7 : Verify directly that $O(n-1, 1)$ is a group, using the defining condition (4.18).

Solution:

The Lorentz group $O(n-1, 1)$ is defined as the set of all $n \times n$ matrices A over the real numbers satisfying the condition $A^T[\eta]A = [\eta]$.

1. Closure: Let $A, B \in O(n-1, 1)$. Then

$$(AB)^T[\eta](AB) = B^T A^T[\eta][A][B] = B^T[\eta]B = [\eta] \implies AB \in O(n-1, 1).$$

2. Associativity: Since $O(n-1, 1)$ is a subgroup of $GL(n, \mathbb{R})$, matrix multiplication in $O(n-1, 1)$ is associative.

3. Existence of an identity: Let I be the identity matrix. Then $I^T[\eta]I = [\eta]$ so $I \in O(n-1, 1)$.

4. Existence of inverses: Let $A \in O(n-1, 1)$. Then

$$[\eta] = I[\eta]I = (AA^{-1})^T[\eta]AA^{-1} = (A^{-1})^T A^T[\eta]AA^{-1} = (A^{-1})^T[\eta]A^{-1}$$

and hence $A^{-1} \in O(n-1, 1)$.

Exercise 4.8 : - Verify that $SU(n)$ and $SO(n)$ are subgroups of $U(n)$ and $O(n)$.

Solution: $SU(n)$ and $SO(n)$ means that the determinant of their respective matrices is equal to 1. For $SU(n)$ we have:

1. Closure: Let $A, B \in SU(n)$. By the properties of determinants, $\det(AB) = \det(A)\det(B) = 1 \cdot 1 = 1$.

2. Associativity: $SU(n)$ automatically inherits associativity from $U(n)$.

3. Existence of an identity: The identity matrix I has determinant equal to 1.

4. Existence of inverses: Let $A \in SU(n)$. Then $\det(A^{-1}) = \frac{1}{\det(A)} = \frac{1}{1} = 1$ and so $A^{-1} \in SU(n)$.

The same logic applies to $SO(n)$, which means that $SU(n)$ is a subgroup of $U(n)$ and $SO(n)$ is a subgroup of $O(n)$.

Exercise 4.10 : Consider an arbitrary matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

and impose the orthogonality condition, as well as $|A| = 1$. Show that (4.22) is the most general solution to these constraints. Then, verify explicitly that $SO(2)$ is a group (even though we already know it is by Exercise 4.8) by showing that the product of two matrices of the form (4.22) is again a matrix of the form (4.22). This will also show that $SO(2)$ is abelian.

Solution: Consider the arbitrary matrix $A \in$ SO(2)

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

.

We have

$$|A| = 1 = ad - bc$$

and

$$A^{-1} = A^T$$

Using these with

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

yields

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

which implies $a = d$ and $-b = c$. Thus

$$|A| = ad - bc = a^2 + c^2 = 1$$

Since $\cos(\theta)$ and $\sin(\theta)$ are both continuously-varying bijective maps from $[0, 1 = 2\pi)$ to $[-1, 1]$, whose respective squares can be made to assume any value in $[0, 1]$ while maintaining a sum of squares equal to 1, there is no loss of generality in parametrizing a and c in terms of θ in this way. Let

$$a = \cos(\theta)$$

$$c = \sin(\theta)$$

which gives

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

which is the most general form and equal to (4.22) from the book.

We now show $SO(2)$ is a group by showing a product of two matrices of this form is again a matrix of this form (i.e. it is automorphic with group operation of matrix multiplication).

Let us now have two matrices A_1 and A_2 defined by

$$A_1 = \begin{pmatrix} \cos(\theta_1) & -\sin(\theta_1) \\ \sin(\theta_1) & \cos(\theta_1) \end{pmatrix}, \; A_2 = \begin{pmatrix} \cos(\theta_2) & -\sin(\theta_2) \\ \sin(\theta_2) & \cos(\theta_2) \end{pmatrix}$$

Then we have

$$A_1 A_2 = \begin{pmatrix} \cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2) & -\cos(\theta_1)\sin(\theta_2) - \sin(\theta_1)\cos(\theta_2) \\ \sin(\theta_1)\cos(\theta_2) + \cos(\theta_1)\sin(\theta_2) & -\sin(\theta_1)\sin(\theta_2) + \cos(\theta_1)\cos(\theta_2) \end{pmatrix}$$

Using the elementary trigonometry identities

$$\cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$$
$$\sin(\alpha+\beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)$$

we obtain

$$A_1 A_2 = \begin{pmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{pmatrix}$$

which is clearly of the general form of $SO(2)$ matrices, and thus this shows that $SO(2)$ forms an Abelian group (it is Abelian since the matrices are commutative).

Exercise 4.11 : Consider an arbitrary complex matrix

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

and impose the unit determinant and unitary conditions. Show that (4.25) is the most general solution to these constraints. Then show that any such solution can also be written in the form (4.26)

Solution: Consider an arbitrary complex matrix

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}.$$

By imposing the unitary condition,

$$A^\dagger A = I$$

and the unit determinant condition

$$|A| = 1$$

we arrive at the following four equations:

$$|\alpha|^2 + |\beta|^2 = 1 \tag{28a}$$
$$|\gamma|^2 + |\delta|^2 = 1 \tag{28b}$$
$$\alpha\overline{\gamma} + \beta\overline{\delta} = 0 \tag{28c}$$
$$\alpha\delta - \beta\gamma = 1. \tag{28d}$$

Solving (28c) and (28d) for β and setting them equal to each other, we see that

$$\begin{aligned} \frac{\alpha\delta - 1}{\gamma} &= -\frac{\alpha\overline{\gamma}}{\overline{\delta}} \\ \Rightarrow |\gamma|^2 &= -|\delta|^2 + \frac{\overline{\delta}}{\alpha}. \end{aligned}$$

Plugging the above result into (28b),

$$\begin{aligned}(-|\delta|^2+\frac{\overline{\delta}}{\alpha})+|\delta|^2 &= 1\\ \Rightarrow \overline{\delta} &= \alpha\\ \Rightarrow \delta &= \overline{\alpha}.\end{aligned}$$

Let us solve (28c) and (28d) for α this time.

$$\begin{aligned}\frac{1+\beta\gamma}{\delta} &= -\frac{\beta\overline{\delta}}{\overline{\gamma}},\\ \Rightarrow |\delta|^2 &= \frac{\overline{\gamma}}{\beta}-|\gamma|^2\end{aligned}$$

and can be plugged into (28b), to get

$$\begin{aligned}|\gamma|^2+(-\frac{\overline{\gamma}}{\beta}-|\gamma|^2) &= 1\\ \Rightarrow \overline{\gamma} &= -\beta.\end{aligned}$$

Complex conjugating both sides, we arrive at the second result:

$$\gamma=-\overline{\beta}.$$

So a generic matrix that satisfies these constraints is of the form

$$\begin{pmatrix}\alpha & \beta\\ -\overline{\beta} & \overline{\alpha}\end{pmatrix}.$$

where $\alpha,\beta\in\mathbb{C}, |\alpha|^2+|\beta|^2=1$. V

Now, consider a generic matrix in $SU(2)$. Since $|\alpha|^2+|\beta|^2=1$, it follows that $|\alpha|,|\beta|\le 1$, always. We may therefore represent any α by the polar form

$$\alpha=e^{i(\psi+\phi)/2}\cos\tfrac{\theta}{2}.$$

Because $|\alpha|^2+|\beta|^2=1$, it must be that β is proportional to $\sin(\frac{\theta}{2})$ since α is proportional to $\cos(\frac{\theta}{2})$. In general, α and β are not necessarily in phase with each other.

$$\beta=ie^{i(\psi-\phi)/2}\sin\tfrac{\theta}{2}$$

for $\phi,\psi\in[0,2\pi],\theta\in[0,\pi]$. Notice the phase difference between α and β. The use of ϕ,ψ, and θ here corresponds to the Euler angles for rotations. This leads to

$$\begin{pmatrix}\alpha & \beta\\ -\overline{\beta} & \overline{\alpha}\end{pmatrix}=\begin{pmatrix}e^{i(\psi+\phi)/2}\cos\frac{\theta}{2} & ie^{i(\psi-\phi)/2}\sin\frac{\theta}{2}\\ ie^{-i(\psi-\phi)/2}\sin\frac{\theta}{2} & e^{-i(\psi+\phi)/2}\cos\frac{\theta}{2}\end{pmatrix}. \tag{29}$$

Thus we have shown that any generic element in $SU(2)$ may also be represented by (29).

Exercise 4.13 : **(a)** Check that L in (4.29) really does represent a boost of velocity β as follows: Use L as a passive transformation to obtain new coordinates (x', y', z', t') from the old ones by

$$\begin{pmatrix}x'\\ y'\\ z'\\ t'\end{pmatrix}=L\begin{pmatrix}x\\ y\\ z\\ t\end{pmatrix}$$

Show that the spatial origin of the unprimed frame, defined by $x=y=z=0$, moves with velocity $-\beta$ in the primed coordinate system, which tells us that the primed coordinate system moves with velocity $+\beta$ with respect to the unprimed system.

(b) A photon traveling in the $+z$ direction has energy-momentum 4-vector given by $(E/c, 0, 0, E)$, where E is the photon energy and we have restored the speed of light c. By applying a boost in the z direction (as given by (4.29)) and using the quantum-mechanical relation $E = h\nu$ between a photon's energy and its frequency ν, derive the ***relativistic doppler shift***

$$\nu' = \sqrt{\frac{1-\beta}{1+\beta}}\,\nu.$$

Solution: **(a)** Plugging in $x = y = z = 0$ into the initial transformation, we have

$$\begin{pmatrix} x' \\ y' \\ z' \\ t' \end{pmatrix} = L \begin{pmatrix} 0 \\ 0 \\ 0 \\ t \end{pmatrix}.$$

This yields

$$\begin{aligned} x' &= -\beta_x \gamma t \\ y' &= -\beta_y \gamma t \\ z' &= -\beta_z \gamma t \\ t' &= \gamma t. \end{aligned}$$

Replacing t with t' for the primed coordinates, we obtain

$$\begin{aligned} x' &= -\beta_x t' \\ y' &= -\beta_y t' \\ z' &= -\beta_z t'. \end{aligned}$$

This shows that the unprimed origin $x = y = z = 0$ is moving away at a velocity of $-\beta$ in the primed frame.

(b) We are given L in equation (4.29):

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & -\beta\gamma \\ 0 & 0 & -\beta\gamma & \gamma \end{pmatrix}.$$

Transforming the null 4-vector $(0, 0, E, E)$ then gives

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & -\beta\gamma \\ 0 & 0 & -\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ E \\ E \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ E' \\ E' \end{pmatrix}.$$

We then have

$$\begin{aligned} E' &= E\gamma - E\beta\gamma \\ &= E\gamma(1-\beta) \\ &= E\frac{1}{\sqrt{1-\beta^2}}(1-\beta) \\ &= E\sqrt{\frac{1-\beta}{1+\beta}} \end{aligned}$$

where we have used $\gamma = \frac{1}{\sqrt{1-\beta^2}}$. Plugging in the quantum-mechanical relation $E = hf$ we come to

$$hf' = hf\sqrt{\frac{1-\beta}{1+\beta}}$$

$$\implies f' = f\sqrt{\frac{1-\beta}{1+\beta}}$$

which is the relativistic doppler shift formula.

Exercise 4.16 : Before moving on to more complicated examples, let's get some practice by acquainting ourselves with a few more basic homomorphisms.

a) First, show that the map

$$\begin{aligned} \exp: \quad & \mathbb{R} \to \mathbb{R}^* \\ & x \mapsto e^x, \end{aligned}$$

from the additive group of real numbers to the multiplicative group of nonzero real numbers, is a homomorphism. Is it an isomorphism? Why or why not?

b) Repeat the analysis from a) for $\exp : \mathbb{C} \to \mathbb{C}^*$.

c) Show that the map

$$\begin{array}{rcl} \det : GL(n, C) & \to & C^* \\ A & \mapsto & \det A \end{array}$$

is a homomorphism for both $C = \mathbb{R}$ and $C = \mathbb{C}$. Is it an isomorphism in either case? Would you expect it to be?

Solution: **(a)** Consider the map

$$\begin{array}{rcl} \exp : \mathbb{R} & \to & \mathbb{R}^* \\ x & \mapsto & e^x. \end{array}$$

Take $a, b \in \mathbb{R}$. Then

$$\exp(a + b) = e^{a+b} = e^a e^b = \exp(a)\exp(b).$$

So indeed exp is a homomorphism from $\{\mathbb{R}, +\} \to \{\mathbb{R}^*, *\}$. However, it is not an isomorphism because it is not surjective. For instance, there is no $x \in \mathbb{R}$ s.t. $\exp(x) = -1$.

(b) Now consider the map exp: $\mathbb{C} \to \mathbb{C}^*$, $x \mapsto e^x$. Take $a, b \in \mathbb{C}$. Then

$$\exp(a + b) = e^{a+b} = e^a e^b = \exp(a)\exp(b).$$

So again exp is a homomorphism, now from $\{\mathbb{C}, +\} \to \{\mathbb{C}^*, *\}$. However, it is not an isomorphism because it is not injective. For instance,

$$\exp(1 + 2\pi i) = e^{1+2\pi i} = e^1 e^{2\pi i} = e^1 = \exp(1)$$

but $1 \neq 1 + 2\pi i$.

(c) Finally, consider the map

$$\begin{array}{rcl} \det : GL(n, C) & \to & C^* \\ A & \mapsto & \det A. \end{array}$$

For either $C = \mathbb{R}$ or $C = \mathbb{C}$:

$$\det(AB) = \det(A)\det(B).$$

So det is certainly a homomorphism. But it is not an isomorphism (unless $n = 1$) because it is not injective. Indeed, any two matrices from SO(n) map to 1 under det, so unless SO(n) contains a single element (as is the case only when $n = 1$), this map is clearly not injective.

Exercise 4.17 : Recall that $U(1)$ is the group of 1×1 unitary matrices. Show that this is just the set of complex numbers z with $|z| = 1$, and that $U(1)$ is isomorphic to $SO(2)$.

Solution: Consider $U(1)$, the set of 1×1 unitary matrices. Clearly, a 1×1 matrix is just a number. Furthermore,

$$a \in U(1) \iff a^\dagger = a^{-1} \iff a^\dagger a = \mathbb{I} \iff \bar{a}a = 1 \iff |a|^2 = 1 \iff |a| = 1.$$

So every element of $U(1)$ contains a single complex number of magnitude 1. But consider that all complex numbers can be written as a magnitude times a phase, and all complex numbers of magnitude one are of the form $z = e^{i\phi}$. Furthermore, we also know a familiar parameterization of SO(2) as rotation by an angle ϕ:

$$A = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}$$

Clearly the mapping

$$f(e^{i\phi}) := \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}$$

is bijective. Is it a homomorphism?

$$f(e^{i\phi}e^{i\theta}) = f(e^{i(\phi+\theta)}) = \begin{pmatrix} \cos(\phi+\theta) & -\sin(\phi+\theta) \\ \sin(\phi+\theta) & \cos(\phi+\theta) \end{pmatrix}$$

$$= \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = f(e^{i\phi})f(e^{i\theta})$$

.

The second to last step in the equality can be verified with double angle formulas. So indeed f is an isomorphism between $U(1)$ and SO(2), meaning that multiplication of complex numbers of magnitude 1 is the same as rotations around the origin in $\mathbb{R}^2$ is some sense.

Exercise 4.18 : Show that the kernel K of any homomorphism $\Phi : G \to H$ is a sub*group* of G. Then determine the kernels of the maps exp and det of Exercise 4.16.

Solution: Consider the homomorphism $\Phi : G \to H$, where G, H are groups. The kernel of Φ is: $K := \{g \in G | \Phi(g) = e_H\}$. Does K satisfy the group axioms?

(1) Closure: Take $k_1, k_2 \in K$. $\Phi(k_1 k_2) = \Phi(k_1)\Phi(k_2) = e_H e_H = e_H$. So $k_1 k_2 \in K$.

(2) Associativity: "Inherited" from G.

(3) Identity: $\Phi(e_G) = e_H$. This is true for any homomorphism. So $e_G \in K$.

(4) Inverses: Take $k \in K$. $\exists k^{-1} \in G$, since G is a group.

$$\Phi(k^{-1}) = \Phi(k^{-1}e) = \Phi(k^{-1}kk^{-1}) = \Phi(k^{-1})\Phi(k)\Phi(k^{-1}) = \Phi(k^{-1})e\Phi(k^{-1}) = \Phi(k^{-1})\Phi(k^{-1}).$$

Since H is a group, we have cancellation, so:

$$e_H = \Phi(k^{-1}).$$

So indeed $k^{-1} \in K$.

Exercise 4.23 : Show that any group, $\mathcal{G}$, with only two elements e and g must be isomorphic to $\mathbb{Z}_2$. To do this, you must define a map Φ: $\mathcal{G} \to \mathbb{Z}_2$ which is one-to-one, onto, and which satisfies equation 4.37. Note that $\mathcal{S}_2$, the symmetric group on two letters, has only two elements. What is the element that corresponds to $-1 \in \mathbb{Z}_2$?

Solution: Consider a group $\mathcal{G}$ consisting of two elements e and g; $\mathcal{G}$ must fulfill the four axioms which define a group. Therefore, $\mathcal{G}$ must contain a unique identity element, which we can take to be e without loss of generality. Since $\mathcal{G}$ must be closed under a multiplication operation, we know $e \cdot g \in \mathcal{G}$, and clearly, $e \cdot g = g$ by definition.

Additionally, because there must exist an inverse to any element in the group, which is also contained within the group, we conclude that g must be its own inverse, or $g = g^-1$, since the only other element in the group is the identity, which is clearly not g's inverse.

We know the elements of $\mathbb{Z}_2 \equiv \{-1, 1\}$, therefore we can now define the map:

$$\Phi : \mathcal{G} \to \mathbb{Z}_2$$

where

$$\Phi(e) = 1$$

and

$$\Phi(g) = -1$$

From this definition, we can immediately see that Φ is one-to-one since it maps exactly one element of $\mathcal{G}$ to exactly one element of $\mathbb{Z}_2$. Also, Φ is onto, since the range of Φ is clearly all of $\mathbb{Z}_2$.

Lastly, we see that $\Phi(e \cdot g) = \Phi(g) = -1 = 1 \cdot -1 = \Phi(e)\Phi(g)$, and Φ satisfies equation 4.36.

Thus, any two element group is isomorphic to $\mathbb{Z}_2$.

As an example, consider the permutation group on two letters $S_2 \equiv \{\sigma_1, \sigma_2\}$, where:

$$\sigma_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

and

$$\sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

An isomorphism from $S_2 \to \mathbb{Z}_2$ would map σ_1, the identity in S_2, to 1, the identity in $\mathbb{Z}_2$, and thus σ_2 would be mapped to -1.

Exercise 4.24 : Let $\mathrm{Diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ be a diagonal matrix whose ith diagonal entry is λ_i. Prove that

$$e^{\mathrm{Diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)} = \mathrm{Diag}\left(e^{\lambda_1}, e^{\lambda_2}, \cdots, e^{\lambda_n}\right)$$

Solution: It is easy to show by induction that $\mathrm{Diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)^k = \mathrm{Diag}\left(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k\right)$

This fact combined with the definition of matrix exponentiation leads us quickly to our goal:

$$\begin{aligned} e^{\mathrm{Diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)} &= \sum_{i=0}^{\infty} \frac{\mathrm{Diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)^k}{k!} \\ &= \sum_{i=0}^{\infty} \frac{\mathrm{Diag}(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k)}{k!} \\ &= \mathrm{Diag}\left(\sum_{i=0}^{\infty} \frac{\lambda_1^k}{k!}, \sum_{i=0}^{\infty} \frac{\lambda_2^k}{k!}, \cdots, \sum_{i=0}^{\infty} \frac{\lambda_n^k}{k!}\right) \\ &= \mathrm{Diag}\left(e^{\lambda_1}, e^{\lambda_2}, \cdots, e^{\lambda_n}\right) \end{aligned}$$

Exercise 4.25 :

a) Show directly from the defining conditions (4.56)–(4.58) that $\mathfrak{so}(n)$, $\mathfrak{u}(n)$, and $\mathfrak{so}(n-1, 1)$ are vector spaces, i.e closed under scalar multiplication and addition. Impose the additional tracelessness condition and show that $\mathfrak{su}(n)$ is a vector space as well.

b) Similarly, show directly from the defining conditions that $\mathfrak{so}(n)$, $\mathfrak{u}(n)$, and $\mathfrak{so}(n-1, 1)$ are closed under commutators.

c) Prove the *cyclic property* of the Trace functional,

$$\mathrm{Tr}(A_1 A_2 \cdots A_n) = \mathrm{Tr}(A_2 \cdots A_n A_1), \quad A_i \in M_n(\mathbb{C})$$

and use this to show directly that $\mathfrak{su}(n)$ is closed under commutators.

Solution:

a) We prove closure under addition and scalar multiplication for $\mathfrak{su}(n)$. The other cases proceed similarly.

To show that $\mathfrak{su}(n)$ is a vector space, let $\Lambda_1, \Lambda_2 \in \mathfrak{su}(n)$. Then Λ_1, Λ_2 are traceless, anti-Hermitian matrices. We need to prove that $\Lambda = \Lambda_1 + \Lambda_2$ is a traceless, anti-Hermitian matrix. To see that Λ is anti-Hermitian, take a look at its i,j-th entry:

$$\Lambda_{(ij)} = \Lambda_{1(ij)} + \Lambda_{2(ij)} = -\bar{\Lambda}_{1(ji)} - \bar{\Lambda}_{2(ji)} = -\bar{\Lambda}_{(ji)}$$

To see that Λ is traceless, we use the linearity of the trace functional:

$$\mathrm{Tr}(\Lambda) = \mathrm{Tr}(\Lambda_1 + \Lambda_2) = \mathrm{Tr}(\Lambda_1) + \mathrm{Tr}(\Lambda_2) = 0$$

b) Solution missing.

c) We now prove the cyclic property of the trace functional. In the problem statment above, if we let $B = A_2 A_3 \cdots A_n$, then our problem is reduced to proving $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

$$\begin{aligned} \mathrm{Tr}(AB) = \mathrm{Tr}(BA) &\Leftrightarrow \mathrm{Tr}(AB) - \mathrm{Tr}(BA) = 0 \\ &\Leftrightarrow \mathrm{Tr}(AB - BA) = 0 \end{aligned}$$

But $AB - BA$ is always traceless for elements of $M_n(\mathbb{C})$: $(AB - BA)_i^j = A_k^i B_j^k - B_k^i A_j^k$, so $\mathrm{Tr}(AB - BA) = (AB - BA)_i^i = A_k^i B_i^k - B_k^i A_i^k = 0$ since we are summing both over both indicies. This proves the cyclic property of the trace functional.

This cyclic property allows us to prove that $\mathfrak{su}(n)$ is closed under brackets. Because $[AB] = AB - BA$ is traceless for *any* elements of $M_n(\mathbb{C})$, it is certainly true of anti-Hermitian, traceless matrices as well. Furthermore,

$$[A, B]^\dagger = (AB - BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB = -[A, B]$$

so $[A, B]$ is anti-hermitian as well. Thus $\mathfrak{su}(n)$ is closed under brackets.

Exercise 4.26 : We want to verify (4.67) by summing the power series for $e^{\theta X}$.

Solution: We want to show $e^{\theta X} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ using
$X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.
The Maclaurin expansion of $e^{\theta X}$ is
$e^{\theta X} = I + \frac{\theta X}{1!} + \frac{(\theta X)^2}{2!} + \frac{(\theta X)^3}{3!} + \frac{(\theta X)^4}{4!} + \frac{(\theta X)^5}{5!} + \frac{(\theta X)^6}{6!} \cdots$
where I is the identity matrix. Thus,
$X^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
$X^3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
$X^4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
and so on. Putting this all together gives
$e^{\theta X} = 1 + \begin{pmatrix} 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} \cdots & -\theta + \frac{\theta^3}{3!} - \frac{\theta^5}{5!} \cdots \\ \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} \cdots & 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} \cdots \end{pmatrix}$

The top left and bottom right element look like the Maclaurin series for $\cos(\theta)$, the top left matches that for $= \sin(\theta)$ and the bottom left that for $\sin(\theta)$

Hence (4.83)

Exercise 4.28 : Using the basis $\mathcal{B} = \{L_x, L_y, L_z\}$ for so(3) show that

$[[X,Y]]_{\mathcal{B}} = [X]_{\mathcal{B}} \times [Y]_B$

The outer brackets signify the column vector associated with the matrices within the brackets. On the Left hand side the inner brackets are showing a standard commutation relation.

Solution: Taking $X = x_1L_x, +y_1L_y + z_1L_z = \begin{pmatrix} 0 & -z_1 & 0 \\ z_1 & 0 & -x_1 \\ -y_1 & x_1 & 0 \end{pmatrix}$

Taking $[Y]_B = (x_2L_x, +y_2L_y + z_2L_z) = \begin{pmatrix} 0 & -z_2 & 0 \\ z_2 & 0 & -x_2 \\ -y_2 & x_2 & 0 \end{pmatrix}$

so(3), these are antisymmetric 3x3 matrices

[X,Y]=XY-YX, using matrix multiplication

$XY = \begin{pmatrix} -(z_1z_2 + y_1y_2) & y_1x_2 & z_1x_2 \\ x_1y_2 & -(z_1z_2 + x_1x_2) & z_1y_2 \\ x_1z_2 & y_1z_2 & -(y_1y_2 + x_1x_2) \end{pmatrix}$

Due to symmetry for YX it is the same but in the subscripts $1 \to 2$ and $2 \to 1$

$XY - YX = \begin{pmatrix} 0 & y_1x_2 - y_2x_1 & x_2z_1 - x_1z_2 \\ x_1y_2 - y_2x_1 & 0 & z_1y_2 - z_2y_1 \\ x_1z_2 - x_2z_1 & y_1z_2 - y_2z_1 & 0 \end{pmatrix}$

This gives the left hand side of the expression we wish to prove

Now as our basis is $\mathcal{B} = \{L_x, L_y, L_z\}$ these are analogous to $\hat{\imath}, \hat{\jmath}\hat{k}$ so in the $\mathcal{B}$ basis

$[X]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$ $[Y]_{\mathcal{B}} = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$

Taking $[X]_B \times [Y]_B = \begin{vmatrix} L_x & L_y & L_z \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} = (y_1z_2 - z_1y_2)L_x + (x_1z_2 - z_1x_2)L_y + (x_1y_2 - y_1x_2)L_z$

These are the elements on the right hand side of the equation we have set out to prove. If you compare these to the element in the matrix formed on the left hand side you can see that they are the same. If you factored out the L_x, L_y, L_z then the two sides would be identical! QED

Exercise 4.31 : Derive the addition law for hyperbolic tangents (4.78) by substituting (4.76) into (4.77).

Solution: Substituting (4.76) into (4.77) yields, after matrix multiplication,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh u_1 \cosh u_2 + \sinh u_1 \sinh u_2 & -\cosh u_1 \sinh u_2 - \sinh u_1 \cosh u_2 \\ 0 & 0 & -\cosh u_1 \sinh u_2 - \sinh u_1 \cosh u_2 & \cosh u_1 \cosh u_2 + \sinh u_1 \sinh u_2 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh(u_1 + u_2) & -\sinh(u_1 + u_2) \\ 0 & 0 & -\sinh(u_1 + u_2) & \cosh(u_1 + u_2) \end{pmatrix}. \qquad (30)$$

We then have

$$\begin{aligned} \tanh(u_1 + u_2) &\equiv \frac{\sinh(u_1 + u_2)}{\cosh(u_1 + u_2)} \\ &= \frac{\sinh u_1 \cosh u_2 + \cosh u_1 \sinh u_2}{\cosh u_1 \cosh u_2 + \sinh u_1 \sinh u_2} \quad \text{by (30)} \\ &= \frac{\tanh u_1 + \tanh u_2}{1 + \tanh u_1 \tanh u_2} \quad \text{by dividing through by } \cosh u_1 \cosh u_2 \end{aligned}$$

as desired.

Exercise 4.40 : Let X, H $\in M_n(C)$. Use (4.94) to show that

$$[X,H]=0 \Longleftrightarrow e^{tX}He^{-tX}=H \quad \forall t\in\mathbb{R}.$$

If we think of H as a quantum-mechanical Hamiltonian, this shows how the invariance properties of the Hamiltonian (like invariance under rotations R) can be formulated in terms of commutators with the corresponding generators.

Solution: Suppose $e^{tX}He^{-tX}=H$, which is the $\Longleftarrow$ part of the claim. Then

$$0=t[X,H]+\frac{t^2}{2}[X,[X,H]]+\frac{t^3}{3}[X,[X,[X,H]]]+\dots$$

which is a polynomial in t. Differentiating this equation and setting $t=0$ then shows that $[X,H]=0$.

The opposite direction of the proposition is trivial and can be shown by simply plugging in $[X,H]=0$.

Exercise 4.41 : Using $[S_i,S_j]=\sum_{k=1}^{3}\epsilon_{ijk}S_k$, compute the matrix representation of ad_{S_i} in the basis B and verify that $[ad_{S_i}]_B=L_i$.

Solution: We know the commutation relation between S_i's, which is $[S_i,S_j]=\sum_{k=1}^{3}\epsilon_{ijk}S_k$. We also know the definition of the linear operator ad_X on a Lie algebra, which is $\mathrm{ad}_X(Y)\equiv[X,Y],\quad X,Y\in\mathfrak{g}$. Then we have the following relations

$$\mathrm{ad}_{S_x}(S_x)=0,\ \mathrm{ad}_{S_x}(S_y)=S_z,\ \mathrm{ad}_{S_x}(S_z)=-S_y$$

$$\mathrm{ad}_{S_y}(S_x)=-S_z,\ \mathrm{ad}_{S_y}(S_y)=0,\ \mathrm{ad}_{S_y}(S_z)=S_x$$

$$\mathrm{ad}_{S_z}(S_x)=S_y,\ \mathrm{ad}_{S_z}(S_y)=-S_x,\ \mathrm{ad}_{S_z}(S_z)=0.$$

We first consider ad_{S_x}. Writing each of the three components in terms of the basis $B=\{S_x,S_y,S_x\}$, we have

$$\mathrm{ad}_{S_x}(S_x)=0\cdot S_x+0\cdot S_y+0\cdot S_z$$

$$\mathrm{ad}_{S_x}(S_y)=0\cdot S_x+0\cdot S_y+1\cdot S_z$$

$$\mathrm{ad}_{S_x}(S_z)=0\cdot S_x+(-1)\cdot S_y+0\cdot S_z.$$

Finally, we identify these three equations as the three column vectors of $[\mathrm{ad}_{S_x}]_B$, and so

$$[\mathrm{ad}_{S_x}]_B=\begin{pmatrix}0&0&0\\0&0&-1\\0&1&0\end{pmatrix}, \tag{31}$$

which is exactly L_x.

Similarly, by carrying out the same procedures above, we can also write both ad_{S_y} and ad_{S_z} in terms of B, obtaining

$$\begin{aligned}[\mathrm{ad}_{S_y}]_B &= \begin{pmatrix}0&0&1\\0&0&0\\-1&0&0\end{pmatrix}\\ [\mathrm{ad}_{S_z}]_B &= \begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}.\end{aligned}$$

Hence, $[\mathrm{ad}_{S_i}]_B=L_i$.

Problems

Problem 4-2 : In this problem, we prove Euler's theorem that any $R \in SO(3)$ has an eigenvector with eigenvalue 1. This means that all vectors v proportional to this eigenvector are invariant under R, i.e. $Rv = v$, and so R fixes a line in space, known as the ***axis of rotation***.

(a) Show that λ being an eigenvalue of R is equivalent to $\det(R - \lambda I) = 0$. Refer to Problem 3-5 if necessary.

(b) Prove Euler's theorem by showing that $\det(R - I) = 0$. Do this using the orthogonality condition and properties of the determinant. You should not have to work in components.

Solution: **Part (a)** If λ is an eigenvalue of R, then we have the relationship $Rv = \lambda v$, where $v \in \mathbb{R}^3$, $v \neq \vec{0}$ is an eigenvector of R. We can rewrite this as:

$$Rv - \lambda v = (R - \lambda I)v = 0.$$

From this, we see that the matrix $(R - \lambda I)$ is not one-to-one, since it sends v to zero, and thus has a nonzero vector in it's kernel. Because $(R - \lambda I)$ is not injective, it cannot be invertible. In problem 3.5, we are asked to prove that a linear operator T is invertible if and only if $|T| \neq 0$. Using this result, we see that $(R - \lambda I)$ is not invertible if and only if its determinant is zero.

Part (b) We know $R \in SO(3)$, so we have $R^T = R^{-1}$ as well as $|R| = 1$. Then:

$$\det(R - I) = \det(R - RR^T) = \det\left(R(I - R^T)\right) = \det(R) \cdot \det(I - R^T) = \det(I - R^T)$$

where we have used the fact that $\det(AB) = \det(A) \cdot \det(B)$.

Knowing also that $(A + B)^T = A^T + B^T$, and that $\det(A) = \det(A^T)$, we can simplify further:

$$\det(I - R^T) = \det(I - R)^T = \det(I - R) = \det\left(-I(-I + R)\right) = \det(-I) \cdot \det(R - I) = -\det(R - I),$$

noting that $\det(-I) = -1$ in $\mathbb{R}^3$.

This leaves us with the expression: $\det(R - I) = -\det(R - I)$, which can only be true if $\det(R - I) = 0$. Thus we have shown that there is always a set of eigenvectors (scalar multiples of each other) in $\mathbb{R}^3$ which is left invariant under a transformation $R \in SO(3)$, and which constitute the line in space we call the *axis of rotation* for that transformation.

Additionally, it is interesting to note as we did in class, that Euler's Theorem only holds for rotations in odd dimensions, since this is when $\det(-I) = -1$. For SO(n) with even n, $\det(-I) = 1$ and the above proof doesn't hold.

Problem 4-3 : Show that the matrix (4.24) is just the component form (in the standard basis) of the linear operator

$$R(\hat{\mathbf{n}}, \theta) = L(\hat{\mathbf{n}}) \otimes \hat{\mathbf{n}} + \cos\theta\,(I - L(\hat{\mathbf{n}}) \otimes \hat{\mathbf{n}}) + \sin\theta\,\hat{\mathbf{n}}\times$$

where (as you should recall) $L(v)(w) = (v|w)$, and the last term eats a vector v and spits out $\sin\theta\,\hat{\mathbf{n}} \times v$. Show that the first term is just the projection onto the axis of rotation $\hat{\mathbf{n}}$, and that the second and third term just give a counterclockwise rotation by θ in the plane perpendicular to $\hat{\mathbf{n}}$. Convince yourself that this is exactly what a rotation about $\hat{\mathbf{n}}$ should do.

Solution: The first term $L(\hat{\mathbf{n}}) \otimes \hat{\mathbf{n}}$ acting on a vector v gives $(\hat{\mathbf{n}}|v)v$, hence is just the projection onto $\hat{\mathbf{n}}$. In components, this operator is just the outer product

$$[L(\hat{\mathbf{n}}) \otimes \hat{\mathbf{n}}] = \begin{pmatrix} n_x n_x & n_x n_y & n_x n_z \\ n_y n_x & n_y n_y & n_y n_z \\ n_z n_x & n_z n_y & n_z n_z \end{pmatrix} \tag{32}$$

As for the second and third terms, they both vanish when applied to $\hat{\mathbf{n}}$, and so only act on the projection of the vector in the plane perpendicular to $\hat{\mathbf{n}}$. The second term multiplies this projection by $\cos\theta$, and the third term adds on a component perpendicular to the original projection (but still in the plane perpendicular to $\hat{\mathbf{n}}$) proportional to $\sin\theta$. This is just what we expect from a rotation in the plane perpendicular to $\hat{\mathbf{n}}$.

In components, the second term is given by $\cos\theta I$ minus $\cos\theta$ times (32). As for the third term, as discussed near the end of Chapter 3 the cross product $\hat{\mathbf{n}}\times$ is given by an antisymmetric tensor, and so the third term is

$$[\sin\theta\hat{\mathbf{n}}\times] = \begin{pmatrix} 0 & -n_z\sin\theta & n_y\sin\theta \\ n_z\sin\theta & 0 & -n_x\sin\theta \\ -n_y\sin\theta & n_x\sin\theta & 0 \end{pmatrix}. \tag{33}$$

Adding these contributions together and re-arranging then yields (4.24)

Problem 4-11 : In this problem we'll calculate the first few terms in the Baker-Campbell-Hausdorff formula (4.66).

Let G be a Lie group and let $X, Y \in \mathfrak{g}$. Suppose that X and Y are small, so that $e^X e^Y$ is close to the identity and hence has a logarithm computable by the power series for ln,

$$\ln(X) = -\sum_{k=1}^{\infty} \frac{(I-X)^k}{k}.$$

By explicitly expanding out the relevant power series, show that up to third order in X and Y,

$$\ln(e^X e^Y) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \cdots \quad .$$

Note that one would actually have to compute higher-order terms to verify that they can be written as commutators, and this gets very tedious. A more sophisticated proof is needed to show that every term in the series is an iterated commutator and hence an element of $\mathfrak{g}$. See Varadarajan [?] for such a proof.

Solution: First note that throughout the problem, we will only be multiplying polynomials in X, Y, and will only keep terms that make a contribution at 3rd order or less.

To third order, then,

$$\begin{aligned} e^X e^Y &= \left(I + X + \frac{X^2}{2} + \frac{X^3}{6} + \cdots\right)\left(I + Y + \frac{Y^2}{2} + \frac{Y3}{6} + \cdots\right) \\ &= I + X + Y + XY + \frac{X^2+Y^2}{2} + \frac{XY^2 + X^2Y}{2} + \frac{X^3+Y^3}{6} + \cdots \end{aligned}$$

We can then use the identity $\ln(X) = -\sum_{k=1}^{\infty} \frac{(I-X)^k}{k}$ to show that

$$
\begin{aligned}
\ln(e^X e^Y) &= -\sum_{k=1}^{\infty} (-1)^k \left(X + Y + XY + \frac{X^2+Y^2}{2} + \frac{XY^2+X^2Y}{2} + \frac{X^3+Y^3}{6} \right)^k /k \\
&= X + Y + XY + \frac{X^2+Y^2}{2} + \frac{XY^2+X^2Y}{2} + \frac{X^3+Y^3}{6} \\
&\quad - \left(X + Y + XY + \frac{X^2+Y^2}{2} \right)^2 /2 \\
&\quad + (X + Y + XY)^3 /3 \\
\\
&= X + Y + XY + \frac{X^2+Y^2}{2} + \frac{XY^2+X^2Y}{2} + \frac{X^3+Y^3}{6} \\
&\quad - \left(X^2 + Y^2 + XY + YX + YXY + XYX + \frac{3(XY^2+X^2Y)}{2} + \frac{YX^2+Y^2X}{2} + X^3 + Y^3 \right) /2 \\
&\quad + \left(XYX + YXY + X^2Y + XY^2 + Y^2X + +YX^2 + X^3 + Y^3 \right) /3 \\
\\
&= X + Y + \frac{1}{2}(XY - YX) - \frac{1}{6}(YXY + XYX) + \frac{1}{12}(X^2Y + XY^2) - \frac{1}{12}(YX^2 + Y^2X) \\
\\
&= X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}(X^2Y - XYX - XYX + YX^2) - \frac{1}{12}(YXY - Y^2X - XY^2 + YXY) \\
\implies \ln(e^X e^Y) &= X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \cdots
\end{aligned}
$$

as desired.

Problem 4-13 : Prove directly that

$$\mathrm{Ad}_{e^{tX}} = e^{t\,\mathrm{ad}_X} \tag{34}$$

by induction, as follows: first verify that the terms first order in t on either side are equal. Then, assume that the nth order terms are equal (where n is an arbitrary integer), and use this to prove that the $n+1$th order terms are equal. Induction then shows that the terms of every order are equal, and so (34) is proven.

Solution: Letting the operators in (34) act on an arbitrary linear operator H, our problem is to show that

$$e^{tX} H e^{-tX} = H + t[X,H] + \frac{t^2}{2!}[X,[X,H]] + \frac{t^3}{3!}[X,[X,[X,H]]] + \dots$$

by induction. The base case is the terms first order in t on either side are equal. After expanding the exponentials e^{tX} and e^{-tX}, we have

$$(1 + tX + \frac{(tX)^2}{2!} + \dots)H(1 - tX + \frac{(tX)^2}{2!} - \dots) = tXH - tHX = t[X,H]$$

So the base case is correct. Now assume that the nth order terms are equal in that equation, which means that

$$t^n \sum_{k=0}^{n} \frac{X^{n-k} H (-X)^k}{(n-k)!k!} = \frac{1}{n!} \mathrm{ad}_{tX}^n (H) \tag{35}$$

We shall show that this equation also holds for $n+1$. To do this, operate on (35) by $t\,\mathrm{ad}X/(n+1)$. The right hand side then becomes $\frac{1}{(n+1)!}\mathrm{ad}_{tX}^{n+1}(H)$, as desired. The left hand side is then

$$t^{n+1} \sum_{k=0}^{n} \frac{X^{n-k+1} H (-X)^k}{(n-k)!k!(n+1)} + \frac{X^{n-k} H (-X)^{k+1}}{(n-k)!k!(n+1)}$$

which can be re-written as

$$t^{n+1}\left[\tfrac{X^{n+1}H}{(n+1)!}+\tfrac{H(-X)^{n+1}}{(n+1)!}+\textstyle\sum_{k=1}^{n}X^{n+1-k}H(-X)^k\left(\tfrac{1}{(n-k)!k!(n+1)}+\tfrac{1}{(n+1-k)!(k-1)!(n+1)}\right)\right]$$
$$=t^{n+1}\left[\tfrac{X^{n+1}H}{(n+1)!}+\tfrac{H(-X)^{n+1}}{(n+1)!}+\textstyle\sum_{k=1}^{n}\tfrac{X^{n+1-k}H(-X)^k}{(n+1-k)!k!}\right]$$
$$=t^{n+1}\textstyle\sum_{k=0}^{n+1}\tfrac{X^{n+1-k}H(-X)^k}{(n+1-k)!k!} \tag{36}$$

which is the left-hand side of (35) but at order $n+1$, concluding the induction step.

Problem 4-14 : In Example (4.44) we claimed that the trace functional Tr is the Lie algebra homomorphism ϕ induced by the determinant function, when the latter is considered as a homomorphism

$$\begin{aligned}\det : GL(n,\mathbb{C}) &\to \mathbb{C}^* \\ A &\mapsto \det A.\end{aligned}$$

This problem asks you to prove this claim.

Begin with the definition (4.87) of ϕ, which in this case says

$$\phi(X)=\frac{d}{dt}\det(e^{tX})|_{t=0}.$$

To show that $\phi(X)=\mathrm{Tr}X$, expand the exponential above to first order in t, plug into the determinant using the formula (3.72), and expand this to first order in t using properties of the determinant. This should yield the desired result.

Solution: To first order in t, we expand $e^X = I+tX$. By the general formula (3.72) for determinants, we have $\det(I+tX)=\epsilon(\overrightarrow{I_1}+t\overrightarrow{X_1},\ldots,\overrightarrow{I_n}+t\overrightarrow{X_n})$ where the $\overrightarrow{X_i}$ are the columns of the matrix X.

Now we can use multilinearity to start expanding this:

$$\begin{aligned}\epsilon(\overrightarrow{I_1}+t\overrightarrow{X_1},\ldots,\overrightarrow{I_n}+t\overrightarrow{X_n}) &= \epsilon(\overrightarrow{I_1},\overrightarrow{I_2}+t\overrightarrow{X_2},\ldots,\overrightarrow{I_n}+t\overrightarrow{X_n})+\epsilon(t\overrightarrow{X_1},\overrightarrow{I_2}+t\overrightarrow{X_2},\ldots,\overrightarrow{I_n}+t\overrightarrow{X_n})\\ &= [\epsilon(\overrightarrow{I_1},\overrightarrow{I_2},\overrightarrow{I_3}+t\overrightarrow{X_3},\ldots,\overrightarrow{I_n}+t\overrightarrow{X_n})+\epsilon(\overrightarrow{I_1},t\overrightarrow{X_2},\overrightarrow{I_3}+t\overrightarrow{X_3},\ldots,\overrightarrow{I_n}+t\overrightarrow{X_n})]\\ &+ [\epsilon(t\overrightarrow{X_1},\overrightarrow{I_2},\overrightarrow{I_3}+t\overrightarrow{X_3},\ldots,\overrightarrow{I_n}+t\overrightarrow{X_n})+\epsilon(t\overrightarrow{X_1},t\overrightarrow{X_2},\overrightarrow{I_3}+t\overrightarrow{X_3},\ldots,\overrightarrow{I_n}+t\overrightarrow{X_n})]\\ &= \ldots\end{aligned}$$

After completely expanding the determinant and keeping only terms first order in t, the RHS becomes

$$\epsilon(\overrightarrow{I_1},\ldots,\overrightarrow{I_n})+\epsilon(t\overrightarrow{X_1},\overrightarrow{I_2},\ldots,\overrightarrow{I_n})+\epsilon(\overrightarrow{I_1},t\overrightarrow{X_2},\ldots,\overrightarrow{I_n})+\ldots+\epsilon(\overrightarrow{I_1},\ldots,\overrightarrow{I_{n-1}},t\overrightarrow{X_n}).$$

We can convert these multilinear forms back to determinants as follows:

$$\det(I)+\det\begin{pmatrix} tX_{11} & 0 & 0 & \ldots & 0\\ tX_{21} & 1 & 0 & \ldots & 0\\ tX_{31} & 0 & 1 & \ldots & 0\\ \vdots & & & \ddots & \vdots\\ tX_{n1} & 0 & 0 & \ldots & 1\end{pmatrix}+\det\begin{pmatrix} 1 & tX_{12} & 0 & \ldots & 0\\ 0 & tX_{22} & 0 & \ldots & 0\\ 0 & tX_{32} & 1 & \ldots & 0\\ \vdots & & & \ddots & \vdots\\ 0 & tX_{n2} & 0 & \ldots & 1\end{pmatrix}+\ldots+\det\begin{pmatrix} 1 & 0 & 0 & \ldots & tX_{1n}\\ 0 & 1 & 0 & \ldots & tX_{2n}\\ 0 & 0 & 1 & \ldots & tX_{3n}\\ \vdots & & & \ddots & \vdots\\ 0 & 0 & 0 & \ldots & tX_{nn}\end{pmatrix}.$$

The determinants of these matrices are easily calculated to be 1, tX_{11}, tX_{22}, ..., and tX_{nn}, respectively. They sum up to $1+\mathrm{Tr}(tX)=1+t\mathrm{Tr}(X)$. Differentiating with respect to t at $t=0$ then gives

$$\phi(X)=\frac{d}{dt}\det(e^{tX})|_{t=0}=\mathrm{Tr}(X)$$

where $\phi(X)$ is the Lie algebra homomorphism induced by the determinant function on $\mathfrak{gl}(n,\mathbb{C})$. Because a continuous homomorphism between matrix Lie groups and the induced Lie algebra homomorphism obey the relation

$$\Phi(e^{tX})=e^{t\phi(X)},$$

and since the determinant is a Lie group homomorphism from $GL(n,\mathbb{C})$ to C, we have

$$e^{\mathrm{Tr} X} = \det e^X$$

as desired.

---

5 Chapter 5

Ch. 5 Exercises

Exercise 5.3 : Verify:

$$[\mathrm{ad}_{\widetilde{L}_i}] = \begin{pmatrix} L_i & 0 \\ 0 & L_i \end{pmatrix} \tag{37}$$

and

$$[\mathrm{ad}_{K_i}] = \begin{pmatrix} 0 & L_i \\ -L_i & 0 \end{pmatrix} \tag{38}$$

Solution: To verify (37), we first compute the components of $\mathrm{ad}_{\tilde{L}_i}$:

$$\begin{aligned}
\mathrm{ad}_{\widetilde{L}_x}(\widetilde{L}_x) &= [\widetilde{L}_x, \widetilde{L}_x] = 0 \\
\mathrm{ad}_{\widetilde{L}_x}(\widetilde{L}_y) &= [\widetilde{L}_x, \widetilde{L}_y] = L_z \\
\mathrm{ad}_{\widetilde{L}_x}(\widetilde{L}_z) &= [\widetilde{L}_x, \widetilde{L}_z] = -L_y \\
\mathrm{ad}_{\widetilde{L}_x}(K_x) &= [\widetilde{L}_x, K_x] = 0 \\
\mathrm{ad}_{\widetilde{L}_x}(K_y) &= [\widetilde{L}_x, K_y] = K_z \\
\mathrm{ad}_{\widetilde{L}_x}(K_z) &= [\widetilde{L}_x, K_z] = -K_y
\end{aligned}$$

Therefore we get:

$$[\mathrm{ad}_{\widetilde{L}_x}] = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} L_x & 0 \\ 0 & L_x \end{pmatrix}$$

Similarly:
$[\mathrm{ad}_{\widetilde{L}_y}] = \begin{pmatrix} L_y & 0 \\ 0 & L_y \end{pmatrix}$
$[\mathrm{ad}_{\widetilde{L}_z}] = \begin{pmatrix} L_z & 0 \\ 0 & L_z \end{pmatrix}$
Therefore:

$$[\mathrm{ad}_{\widetilde{L}_i}] = \begin{pmatrix} L_i & 0 \\ 0 & L_i \end{pmatrix}$$

To verify (38), we proceed similarly:

$\mathrm{ad}_{K_x}(K_x) = [K_x, K_x] = 0$
$\mathrm{ad}_{K_x}(K_y) = [K_x, K_y] = -\widetilde{L}_z$
$\mathrm{ad}_{K_x}(K_z) = [K_x, K_z] = \widetilde{L}_y$
$\mathrm{ad}_{K_x}(\widetilde{L}_x) = [K_x, \widetilde{L}_x] = 0$
$\mathrm{ad}_{K_x}(\widetilde{L}_y) = [K_x, \widetilde{L}_y] = K_z$
$\mathrm{ad}_{K_x}(\widetilde{L}_z) = [K_x, \widetilde{L}_z] = -K_y$

Therefore:

$$[\mathrm{ad}_{K_x}] = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -L_x \\ L_x & 0 \end{pmatrix}$$

Similarly: $[\mathrm{ad}_{K_y}] = \begin{pmatrix} 0 & -L_y \\ L_y & 0 \end{pmatrix}$; $[\mathrm{ad}_{K_z}] = \begin{pmatrix} 0 & -L_z \\ L_z & 0 \end{pmatrix}$

Therefore:

$$[\mathrm{ad}_{K_i}] = \begin{pmatrix} 0 & -L_i \\ L_i & 0 \end{pmatrix}$$

Exercise 5.5 : Use the definition of induced Lie algebra representations and the explicit form of e^{tS_i} in the fundamental representation (which can be deduced from (4.74) to compute the action of the operators $\pi_1(S_i)$ on the functions z_1 and z_2 in $P_1(\mathbb{C}^2)$. Show that we can write these operators in differential form as

$$\begin{aligned} \pi_1(S_1) &= \frac{i}{2}\left(z_2\frac{\partial}{\partial}z_1 + z_1\frac{\partial}{\partial}z_2\right) \\ \pi_1(S_2) &= \frac{1}{2}\left(z_2\frac{\partial}{\partial}z_1 - z_1\frac{\partial}{\partial}z_2\right) \\ \pi_1(S_3) &= \frac{i}{2}\left(z_1\frac{\partial}{\partial}z_1 - z_2\frac{\partial}{\partial}z_2\right). \end{aligned} \tag{39}$$

Prove that these expressions also hold for the operators $\pi_l(S_i)$ on $P_l(\mathbb{C}^2)$. Verify the $\mathfrak{su}(2)$ commutation relations directly from these expressions. Finally, use (5.12) to show that

$$(\pi_l(S_z))(z_1^{l-k}z_2^k) = i(l/2-k)z_1^{l-k}z_2^k.$$

Warning: This is a challenging exercise, but worthwhile. See the next example for an analogous calculation for $SO(3)$. Also, don't forget that for a Lie group representation $\mathit{\Pi}$, induced Lie algebra representation π, and Lie algebra element X,

$$\pi(X) = \frac{d}{dt}\mathit{\Pi}(e^{tX}) \neq \mathit{\Pi}\left(\frac{d}{dt}e^{tX}\right),$$

as discussed in Box 4.7.

Solution: From the definition of the induced Lie algebra we have:

$$\pi_1(S_i) = \frac{d}{dt}\Pi(e^{tS_i})|_{t=0} \tag{40}$$

Giving:

$$(\pi_1(S_i)p)(\vec{v}) = \frac{d}{dt}(p(e^{-tS_i})\vec{v}) = p(-S_i\vec{v}) \tag{41}$$

Since:

$$-S_1\vec{v} = \frac{i}{2}\begin{pmatrix} z_2 \\ z_1 \end{pmatrix} \tag{42}$$

$$-S_2\vec{v} = \frac{1}{2}\begin{pmatrix} z_2 \\ -z_1 \end{pmatrix} \tag{43}$$

$$-S_3\vec{v} = \frac{i}{2}\begin{pmatrix} z_1 \\ -z_2 \end{pmatrix} \tag{44}$$

We have from ??:

$$\begin{aligned}
\pi_1(S_1)z_1 &= \frac{i}{2}z_2 \\
\pi_1(S_1)z_2 &= \frac{i}{2}z_1 \\
\pi_1(S_2)z_1 &= \frac{1}{2}z_2 \\
\pi_1(S_2)z_2 &= -\frac{1}{2}z_1 \\
\pi_1(S_3)z_1 &= \frac{i}{2}z_1 \\
\pi_1(S_3)z_2 &= -\frac{i}{2}z_2
\end{aligned} \tag{45}$$

Therefore we clearly have,

$$\begin{aligned}
\pi_1(S_1)z_1 &= \frac{i}{2}(z_2\frac{\partial}{\partial z_1} + z_1\frac{\partial}{\partial z_2})z_1 \\
\pi_1(S_2)z_1 &= \frac{1}{2}(z_2\frac{\partial}{\partial z_1} - z_1\frac{\partial}{\partial z_2})z_1 \\
\pi_1(S_3)z_1 &= \frac{i}{2}(z_1\frac{\partial}{\partial z_1} - z_2\frac{\partial}{\partial z_2})z_1
\end{aligned}$$

And

$$\begin{aligned}
\pi_1(S_1)z_2 &= \frac{i}{2}(z_2\frac{\partial}{\partial z_1} + z_1\frac{\partial}{\partial z_2})z_2 \\
\pi_1(S_2)z_2 &= \frac{1}{2}(z_2\frac{\partial}{\partial z_1} - z_1\frac{\partial}{\partial z_2})z_2 \\
\pi_1(S_3)z_2 &= \frac{i}{2}(z_1\frac{\partial}{\partial z_1} - z_2\frac{\partial}{\partial z_2})z_2
\end{aligned}$$

Showing that:

$$\begin{aligned}
\pi_1(S_1) &= \frac{i}{2}(z_2\frac{\partial}{\partial z_1} + z_1\frac{\partial}{\partial z_2}) \\
\pi_1(S_2) &= \frac{1}{2}(z_2\frac{\partial}{\partial z_1} - z_1\frac{\partial}{\partial z_2}) \\
\pi_1(S_3) &= \frac{i}{2}(z_1\frac{\partial}{\partial z_1} - z_2\frac{\partial}{\partial z_2})
\end{aligned}$$

Part 2:
Prove that these expressions also hold for the operators $\pi_l(S_i)$ on $P_l(C^2)$:

$$\begin{aligned}
(\pi_l(S_i)p)\vec{v} &= \frac{d}{dt}(\Pi e^{tS_i}p)\vec{v} \\
&= \frac{d}{dt}p(e^{-tS_i}\vec{v}) \\
&= \sum_{j=1}^{2}\frac{\partial p}{\partial v^j}(-S_i\vec{v})^j
\end{aligned} \tag{46}$$

Eqn. (46) comes from the multivariable chain rule. Hence we have:

$$\pi(S_i) = \sum_{j=1}^{2}(-S_i\vec{v})^j\frac{\partial}{\partial v^j} \tag{47}$$

Knowing that $v^i = z_i$ and using the previous equations we have

$$\sum_{j=1}^{2}(-S_1\vec{v})^j \frac{\partial}{\partial v^j} = \frac{i}{2}\sum_{j=1}^{2}\begin{pmatrix} z_2 \\ z_1 \end{pmatrix}^j \frac{\partial}{\partial v^j} = \frac{i}{2}(z_1\frac{\partial}{\partial z_2} + z_2\frac{\partial}{\partial z_1}) \ . \tag{48}$$

Therefore find that (46) is equivalent to:

$$\begin{aligned}
\pi(S_1) &= \frac{i}{2}(z_2\frac{\partial}{\partial z_1} + z_1\frac{\partial}{\partial z_2}) \\
\pi(S_2) &= \frac{1}{2}(z_2\frac{\partial}{\partial z_1} - z_1\frac{\partial}{\partial z_2}) \\
\pi(S_3) &= \frac{i}{2}(z_1\frac{\partial}{\partial z_1} - z_2\frac{\partial}{\partial z_2})
\end{aligned}$$

Part 3:
Show that $(\pi_l(S_z))(z_1^{l-k}z_2^k) = i(l/2-k)z_1^{l-k}z_2^k$ using (5.12):

$$\begin{aligned}
(\pi_l(S_z))(z_1^{l-k}z_2^k) &= \frac{i}{2}(z_1\frac{\partial}{\partial z_1} - z_2\frac{\partial}{\partial z_1})(z_1^{l-k}z_2^k) \\
&= \frac{i}{2}(z_1(l-k)z_1^{l-k-1}z_2^k - z_1^{l-k}z_2 k z_2^{k-1}) \\
&= \frac{i}{2}z_1^{l-k}z_2^k(l-2k) = i(l/2-k)z_1^{l-k}z_2^k
\end{aligned}$$

as desired.

Exercise 5.11 : Verify that (5.21) defines a Lie algebra representation. Mainly, this consists of verifying that

$$[(\pi_1 \otimes \pi_2)(X), (\pi_1 \otimes \pi_2)(Y)] = (\pi_1 \otimes \pi_2)([X, Y]).$$

Solution: If (5.21) gives a Lie algebra representation it must preserve the Lie bracket:

$$[(\pi_1 \otimes \pi_2)(X), (\pi_1 \otimes \pi_2)(Y)] = (\pi_1 \otimes \pi_2)([X, Y]).$$

We show this by explicitly expanding the commutator:

$$[(\pi_1 \otimes \pi_2)(X), (\pi_1 \otimes \pi_2)(Y)] = [\pi_1(X) \otimes I + I \otimes \pi_2(X), \pi_1(Y) \otimes I + I \otimes \pi_2(Y)]$$

when we expand this further, we see we only get two nonzero terms (terms like $I \otimes \pi_2(X)$ and $\pi_1(Y) \otimes I$ commute):

$$\begin{aligned}
[(\pi_1 \otimes \pi_2)(X), (\pi_1 \otimes \pi_2)(Y)] &= [\pi_1(X) \otimes I, \pi_1(Y) \otimes I] + [I \otimes \pi_2(X), I \otimes \pi_2(Y)] \\
&= (\pi_1(X)\pi_1(Y) - \pi_1(Y)\pi_1(X)) \otimes I \\
&\quad + I \otimes (\pi_2(X)\pi_2(Y) - \pi_2(Y)\pi_2(X)) \\
&= [\pi_1(X), \pi_1(Y)] \otimes I + I \otimes [\pi_2(X), \pi_2(Y)]
\end{aligned}$$

Because π_1 and π_2 are Lie algebra homomorphisms then for all elements X, Y of the Lie algebra:

$$[\pi_1(X), \pi_1(Y)] = \pi_1([X, Y])$$

and

$$[\pi_2(X), \pi_2(Y)] = \pi_2([X, Y]).$$

Then if this is true:

$$\begin{aligned}
[(\pi_1 \otimes \pi_2)(X), (\pi_1 \otimes \pi_2)(Y)] &= \pi_1([X, Y]) \otimes I + I \otimes \pi_2([X, Y]) \\
&= (\pi_1 \otimes \pi_2)([X, Y])
\end{aligned}$$

as desired.

Exercise 5.15 : Prove (5.31). As in Exercise 5.14, this can be done by either computing the infinitesimal form of (5.29), or performing a calculation similar to the one above (5.29), starting with the Lie algebra representation associated with (5.28)

Solution: The Lie algebra representation induced by the (1,1) tensor representation is:

$$\begin{aligned}\pi_1^1(X)\,T &\equiv \frac{d}{dt}\left(\Pi_1^1\left(e^{tX}\right)T\right)_{t=0}\\ &= \frac{d}{dt}\left(\Pi\left(e^{tX}\right)T\Pi^{-1}\left(e^{tX}\right)\right)_{t=0}\end{aligned}$$

We also know the following is true:

$$\Pi^{-1}\left(e^{tX}\right)=\Pi\left(\left(e^{tX}\right)^{-1}\right)=\Pi\left(e^{-tX}\right).$$

and so

$$\begin{aligned}\pi_1^1(X)\,T &= \frac{d}{dt}\left(\Pi\left(e^{tX}\right)T\Pi\left(e^{-tX}\right)\right)_{t=0}\\ &= \frac{d}{dt}\left(\Pi\left(e^{tX}\right)\right)T\Pi\left(e^{-tX}\right)+\Pi\left(e^{tX}\right)T\frac{d}{dt}\left(\Pi\left(e^{-tX}\right)\right)|_{t=0}\\ &= \pi(X)\,T+T\pi(-X)\\ &= \pi(X)\,T-T\pi(X)\\ &= [\pi(X)\,,T]\end{aligned}$$

as desired.

Exercise 5.30 : Show that if $V=\bigoplus_{i=1}^{k}W_i$, then $W_i\cap W_j=\{0\}\ \forall\, i\neq j$, i.e. the intersection of two different W_i is just the zero vector. Verify this explicitly in the case of the two decompositions in (5.52).

Solution: If $v\in(V=W_i+W_j)$ then $v=a+b$ for some $a\in W_i$ and $b\in W_j$ and more generally, $v=(a+c)+(b-c)$ for any such c that $c\in W_i$ and $c\in W_j$. That is to say, we have $c\in W_i\cap W_j$. But, if $v=a+b$ is a unique representation, then c must be the zero vector, hence $W_i\cap W_j=\{0\}$.

Exercise 5.31 : Show that $V=\bigoplus_{i=1}^{k}W_i$ is equivalent to the statement that the set

$$\mathcal{B}=\mathcal{B}_1\cup\cdots\cup\mathcal{B}_k,$$

where each $\mathcal{B}_i$ is an arbitrary basis for W_i, is a basis for V.

Solution: We have $v=\sum_{i=1}^{k}w_i$ for any $v\in V$, so the bases must span V, i.e. $w_i=\sum_j c_ie_{ij}$ where

$$\mathcal{B}_i=\{e_{ij}\}_{j=1,\ldots,\dim w}$$

Suppose $v=\sum_i\sum_j c_i^je_{ij}=\sum_i\sum_j d_i^je_{ij}$

Then $v=\sum_i\sum_j(c_i^j-d_i^j)e_{ij}=0$

Then $c_i^j-d_i^j=0$ due to uniqueness. Thus we must have linear independence. Conversely, it is clear that $\mathcal{B}$ spans V and it immediately follows that due to the linear independence of $\mathcal{B}_i$ each $v\in V$ can be written uniquely, thus the two statements are equivalent.

Ch. 5 Problems

Problem 5-8 : In this problem we'll meet a representation that is *not* completely reducible.

a) Consider the representation $(\mathit{\Pi}, \mathbb{R}^2)$ of $\mathbb{R}$ given by

$$\begin{aligned} \mathit{\Pi} : \mathbb{R} &\rightarrow GL(2, \mathbb{R}) \\ a &\mapsto \mathit{\Pi}(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Verify that $(\mathit{\Pi}, \mathbb{R}^2)$ is a representation.

b) If $\mathit{\Pi}$ was completely reducible then we'd be able to decompose $\mathbb{R}^2$ as $\mathbb{R}^2 = V \oplus W$ where V and W are one-dimensional. Show that such a decomposition is impossible.

Solution:

(a) We have

$$\Pi(a)\Pi(b) = \begin{pmatrix} 1 & a+b \\ 0 & 1 \end{pmatrix} = \Pi(a+b)$$

so $(\Pi, \mathbb{R}^2)$ is a representation.

(b) We have

$$\Pi(a) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + ay \\ y \end{pmatrix} . \tag{49}$$

So we have that the nontrivial, invariant subspace $y = 0$ is fixed, hence $(\Pi, \mathbb{R}^2)$ is not irreducible. But, there is no complementary invariant subspace. If there was, it would be 1-D (i.e an eigenspace) and its nonzero vectors must have $y \neq 0$. In this case these must then be eigenvectors with eigenvalue 1 (by (49), but then we must have $x = x + ay$ $\forall a$, which implies $y = 0$, a contradiction. Thus, there can be no complementary invariant subspace.

References

[Jee15] Nadir Jeevanjee. *An Introduction to Tensors and Group Theory for Physicists*. Springer International Publishing, New York, NY, 2nd edition, 2015.