Chapter 13. Symmetry Groups

Groups

Definitions:
A **group** is a set G with an operation $\circ$ that is closed and associative, has an identity e, and every element g has an inverse g^{-1} such that $g \circ g^{-1}$ = e = $g^{-1} \circ g$.

A group G is **Abelian** if it is commutative: $g \circ h$ = $h \circ g$ for all g, h in G.

Often the operation for an Abelian group is denoted " + " and for a non-Abelian group " $*$ " or " $\times$ ".

A **subgroup** is a subset of G that is a group under $\circ$.

Let H be a subgroup of G. A **right coset of H** is a set H$\circ g$ = {$h \circ$ g: $h \in$H}, where $g \in G$. A **left coset** is $g \circ \mathrm{H}$.The only coset that is a group is the set H itself: $\mathrm{H} = \mathrm{H} \circ e = e \circ \mathrm{H}$ where e is the identity element of G. The cosets of H form a **partition of G**, the union of disjoint sets: $\mathrm{G} = \bigcup_{g \in \mathrm{G}} \left(\mathrm{H} \circ g\right)$. If G is Abelian, the left and right cosets are identical.

A **normal subgroup** is a subgroup H that satisfies $g \circ$H = H$\circ g$ for all g in G, or equivalently $\mathrm{H} = g^{-1}\mathrm{H} \circ g$. Thus every subgroup of an Abelian group is normal.

A group is **simple** if it contains no non-trivial normal subgroup. The simple groups are the fundamental “building blocks” of more complex groups.

Theorem. The simple finite groups have been classified into classical and exceptional groups. The largest exceptional group has ≈ 10^{60} elements and is known as **the monster.**

Definition. The **Product Group** of G and H is **GxH** = $\left\{\left(g,h\right): g \in G, h \in H\right\}$ with group operation $\left(g_1,h_1\right) \circ \left(g_2,h_2\right) = \left(g_1 \circ g_2, h_1 \circ h_2\right)$.

Definition. Let N be a subgroup of G. The **Factor Space G/N** or $\dfrac{\mathbf{G}}{\mathbf{N}}$ is the collection of cosets N$\circ g$ along with the $*$ operation:

$$\left(\mathrm{N} \circ g_1\right) * \left(\mathrm{N} \circ g_2\right) = \mathrm{N} \circ \left(g_1 g_2\right).$$

Theorem. If N is normal then G/N is a group, called the **Factor Group.**

Definition. Two groups G and $\bar{\mathrm{G}}$ are **isomorphic** if there is a 1-1 map $f: \mathrm{G} \to \bar{\mathrm{G}}$ from G onto $\bar{\mathrm{G}}$ that preserves the group operation:

$$f(g_1 \circ g_2) = f(g_1) \circ f(g_2) \equiv \bar{g}_1 \circ \bar{g}_2 .$$

We denote this by $\bar{\mathbf{G}} \cong \mathbf{G}$.

Theorem. [13.10] $\mathrm{H} \cong \dfrac{\mathrm{G} \times \mathrm{H}}{\mathrm{G}}$.

Note that the group operation is a function, $\circ: \mathrm{G} \times \mathrm{G} \to \mathrm{G}: \circ(g_1, g_2) \equiv g_1 \circ g_2$. If G is also a topological space (i.e., a set with a topology, a structure of open and closed sets), then $\circ$ can either be continuous or not.

Definition. A **group (G, $\circ$) is continuous** if both $\circ: \mathrm{G} \times \mathrm{G} \to \mathrm{G}$ and the inversion operation are continuous functions when G is considered as a topological space. A **Lie group** is a continuous group that is locally homeomorphic to $\mathbb{R}^n$ in which a differentiation structure has been defined.

Theorem. There are precisely 4 **classical** and 5 **exceptional** simple Lie groups.

- Classical Families: A_m, B_m, C_m, D_m having dimensions $m(m+2)$, $m(2m+1)$, $m(2m+1)$, and $m(2m-1)$, respectively where $m \in \mathbb{Z}^+$.
- Exceptional Groups: E_6, E_7, E_8, F_4, G_2 of dimension 78, 133, 248, 52, and 14 respectively

The **dimension of a group** is its dimension as a topological space, which we now define. Intuitively the dimension of a space is 1 + dim (boundary of space). For example, a disk has dimension 2 and its boundary, a circle, has dimension 1. A line segment has dimension 1 and its boundary, 2 points, has dimension zero. Since a point has dimension 0, its boundary, the empty set $\varnothing$, is defined to have dimension -1. It is standard to start with $\varnothing$ and define topological dimension inductively.

There are several definitions of dimension. We give the **Small Inductive Dimension** and briefly two others.

Definition. A collection $\mathcal{U} = \{U_\alpha\}$ is an **open basis for a topological space** X if every open set is a union of sets from $\mathcal{U}$. We first give a few preliminaries.

Examples. In $\mathbb{R}$, the collection of open intervals forms a basis. In $\mathbb{R}^3$, the collection of open balls forms a basis.

Definition. The **closure** $\overline{A}$ of a set A is the smallest closed set that contains A. The **boundary** of A is $\partial A = \overline{A} - A$.

Definition. Let X be a topological space and $\mathcal{U}$ an open basis for X.

(1) **dim X = -1** if $X = \varnothing$

(2) dim X ≤ n if for all points x and open sets W such that $x \in W$ there exists $U \in \mathcal{U}$ such that $x \in U \subseteq \overline{U} \subseteq W$ and dim $\partial U \leq n - 1$

(3) **dim X = *n*** if (2) is true for n but false for $n - 1$

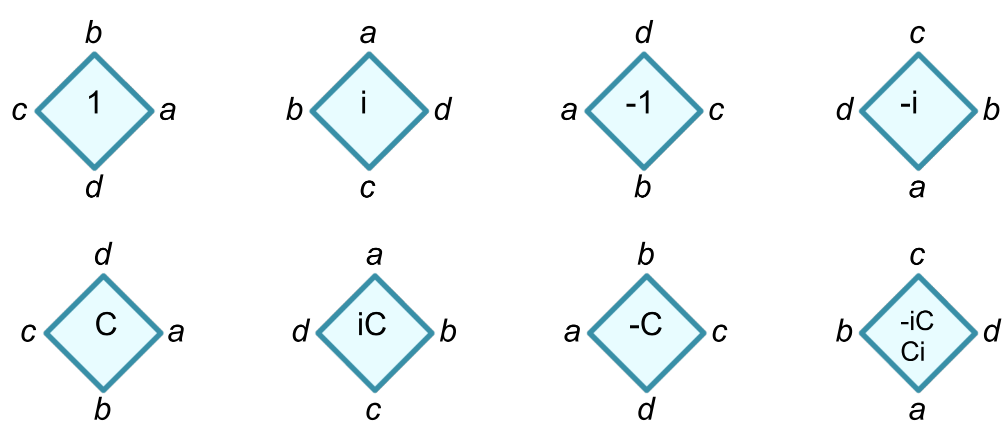
(4) **dim X =** ∞ if for every n, dim X ≤ n is false

The **Large Inductive Dimension** is similar but replaces points with closed sets.

The **Lebesgue Covering Dimension** defines dim X = n if every open cover of X has an open refinement in which no point belongs to more that n + 1 sets in the refinement.

All of these definitions agree on spaces like $\mathbb{R}^n$ that are separable and metrizable.

Symmetries of a Square



Definitions:

- **Non-reflecting Group:** < i > = {1, i, -1, -i}
- **Reflecting Group:** < i, C > = {1, i, -1, -i, C, iC, -C, -iC = Ci}
- **C** is complex conjugation: $a + bi \mapsto a - bi$. **1** is the null rotation, which is the group identity element. ***i*** is the 90° counter-clockwise rotation of the square

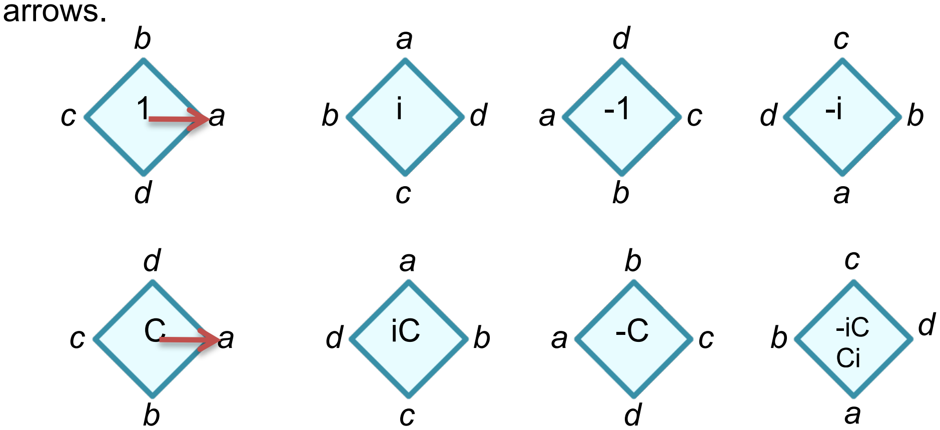
Convention: ab means b acts first.

Definition. A subgroup of a symmetry group is called a **reduced symmetry group**.

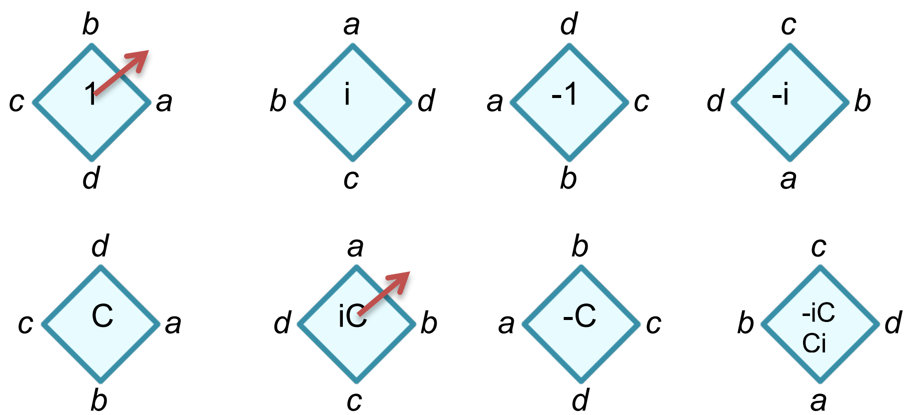
Examples:

- Normal subgroups of < i, C > :
  - {1, -1, C, -C} , {1, -1}, {1, -1}
- Non-normal subgroups of < i, C > :
  - {1, -C}, {1, iC}, {1, C}
    - For example, {1, C} i = {i, Ci} ≠ {i, -Ci} = i {1, C}

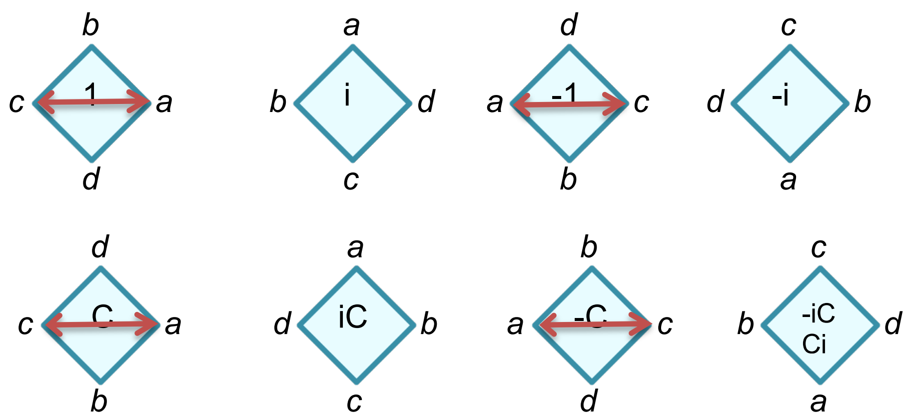
Example [13.6]: Reduced symmetry groups can be generated using one or more arrows.



{1, C} is a reduced symmetry group



{1, iC} is a reduced symmetry group



{1, -1, C, -C} is a reduced symmetry group

Symmetries of a Sphere

Definitions:

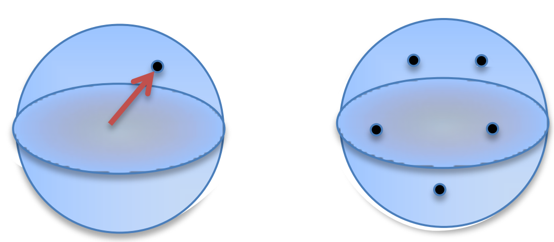
- **SO(3)** is the group of non-reflective symmetries of a 3-sphere
- **O(3)** is the **Orthogonal Group**. It consists of both the reflective and non-reflective symmetries of a sphere.
- O(3) = SO(3) ∪ T, the disjoint union of O(3) with the coset of reflective symmetries
- $\mathbf{T} = \mathrm{R}\,\mathrm{SO}(3) = \{\mathrm{R}g : g \in \mathrm{SO}(3)\}$ where **R** is the reflection operator on the sphere.

Recall problem [12.7]: SO(3) is group isomorphic to the solid sphere $\mathcal{R}$ of radius π with antipodal points identified.

Theorem. (Problem [13.7]) SO(3) and {1, R} are the only normal subgroups of O(3), where **1** is the null rotation. (Penrose overlooked that the latter group is normal.)

Examples. Reduced Symmetry Groups

The set of rotations that fix a point on the sphere forms a non-normal subgroup. It is the set of rotations having the arrow as its axis.



Marking the sphere with vertices of a regular polyhedron reduces to the finite group of rotations of the sphere that take each vertex to one of the others. Such reduced symmetry groups are non-normal.

Linear Transformations and Matrices

Definition. A nonempty set (R, +, •) is a **ring** if for all *a*, *b*, *c* in R:

(1) (R, +) is an Abelian group
(2) R is closed under multiplication •
(3) $a \bullet (b + c) = a \bullet b + a \bullet c$ and $(b + c) \bullet a = b \bullet a + c \bullet a$ (left and right distributive)

A ring R is an **associative ring** if it is associative under multiplication:

(4) $r \bullet (s \bullet t) = (r \bullet s) \bullet t$ for $r, s, t \in \mathrm{R}$

There are rings that have no multiplicative identity (i.e., no element 1). Rings that do have a multiplicative identity are said to be **rings with unit element**.

Definition. A **field** is a ring F where the non-zero elements form an Abelian group under multiplication.

Definition. (V,+) is a **vector space over a field F** if (V,+) is an Abelian group and $\forall\, \alpha, \beta \in \mathrm{F}$ and $v, w \in V$:

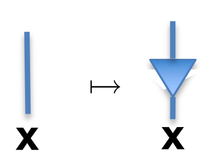
(1) $\alpha(v + w) = \alpha v + \alpha w$
(2) $(\alpha + \beta)v = \alpha v + \beta v$
(3) $\alpha(\beta v) = (\alpha\beta)v$
(4) $1v = v$

The elements of F are called **scalars.** The elements of V are called **vectors.**

Definition. Let V and W be vector spaces.

- $f : \mathrm{V} \to \mathrm{W}$ is a **homomorphism** if *it* preserves the vector space structure:
  - $f(au + bv) = af(u) + bf(v)$ for all vectors *u* and *v* and scalars *a* and *b*.
- **Hom(V,W)** is the set of homomorphisms from V to W.
- $\mathcal{A}$**(V)** = Hom(V,V).
- A **linear transformation** is a member $T \in \mathcal{A}(\mathrm{V})$.
  - That is, a linear transformation is a function $T : \mathrm{V} \to V$ such that $T(au + bv) = aTu + bTv$.

Theorem. [13.12, 13.13] Let V = $\mathbb{R}^3$, using (x^1, x^2, x^3) instead of (*x*, *y*, *z*). Then a linear transformation *T* on V takes the form $T : x^r \mapsto T^r{}_s x^s = ax^1 + bx^2 + cx^3$ where $a, b, c \in \mathbb{R}$. In diagrammatic form this is



Proof. Linear transformations are represented by matrices:

$$T: \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \mapsto \begin{pmatrix} T^1{}_1 & T^1{}_2 & T^1{}_3 \\ T^2{}_1 & T^2{}_2 & T^2{}_3 \\ T^3{}_1 & T^3{}_2 & T^3{}_3 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \quad \text{where } T^i{}_j \in \mathbb{R},$$

or

$$\mathrm{x} \mapsto Tx,$$

or

$$x^r \mapsto T^r{}_s x^s = T^r{}_1 x^1 + T^r{}_2 x^2 + T^r{}_3 x^3 \ . \qquad \blacksquare$$

Definition. The **transpose** of the matrix $T = T^i{}_j$ is the matrix $T^{\mathsf{T}} = \left(T^{\mathsf{T}}\right)^i{}_j = T^j{}_i$.

Note. The indices are dummy variables. We can express T as $T = T^r{}_s$ and also $T = T^s{}_r$. This can become subtle when expressing either the transpose or the inverse of T. One cannot always write $T = T^s{}_r$ for the transpose. The transpose must be expressed as $\left(T^{\mathsf{T}}\right)^r{}_s$ or $\left(T^{\mathsf{T}}\right)^s{}_r$ unless it is part of an expression involving T such as $T^s{}_r = \left(T^r{}_s\right)^{\mathsf{T}}$.

Definition. The **inverse** of the matrix T is the matrix T^{-1} satisfying $TT^{-1} = I = T^{-1}T$. If $S = T^{-1}$, we can also write $S^r{}_s = \left(T^r{}_s\right)^{-1}$ and $S^s{}_r = \left(T^s{}_r\right)^{-1}$ but not $S^s{}_r = \left(T^r{}_s\right)^{-1}$. The indices should match.

Definition. A matrix T is **orthogonal** if $T^{-1} = T^{\mathsf{T}}$.

Theorem. If $R = ST$ then $R^a{}_c = S^a{}_b T^b{}_c$. That is, the composition, R, of 2 linear transformations is the result of matrix multiplication of S and T. In diagrammatic notation:

R =  =  = ST

Example. $TI = T = IT$ is written in diagrammatic form as  =  = .

In $\mathbb{R}^3$, $I = \delta^a_b = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if a} \neq \text{b} \end{cases} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ where *a*, *b* range over {1, 2, 3}.

Definitions. A linear transformation *T* is **singular** if Dim(*T*V) < Dim W; that is, *T* is not *onto*.

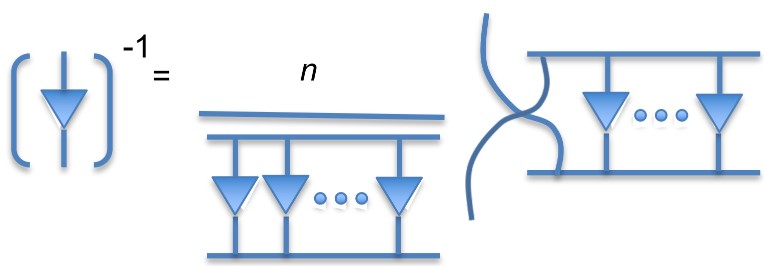
Theorem. [13.17] *T* is singular iff $\exists\ v \neq 0$ such that $Tv = 0$.

Corollary. [Bud] *T* is 1-1 iff *T* is non-singular iff *T* is onto.

Proof. *T* is 1-1 $\Leftrightarrow \forall v \neq w\ \ T(v-w) = T(v) - T(w) \neq 0 \overset{(*)}{\Leftrightarrow} \forall u \neq 0\ T(u) \neq 0$

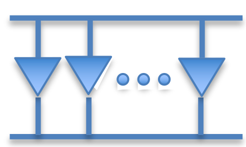
$\overset{[13.17]}{\Leftrightarrow}$ *T* is non-singular $\Leftrightarrow$ *T* is onto.
(*) Set $v = 3u$ and $w = 2u$. ■

Theorem. [13.18] If *T* is nonsingular, then it has an inverse T^{-1}.

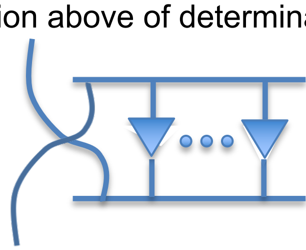
Theorem. [13.19] T^{-1} = 

if *T* is non-singular.

Determinants and Traces

Definition. Det $T = \dfrac{1}{n!}$  $= \dfrac{1}{n!} \in^{ab\cdots d} T^e{}_a T^f{}_b \cdots T^h{}_d\, \varepsilon_{ef\cdots h}$.

Note: Since $T^{-1} = \dfrac{\text{Adj}(T)}{\text{Det } T}$ (**Adjugate Formula)**, from Theorem 13.19 and the definition above of determinant we get that

 = $(n-1)!$ Adj (T),

where Adj (*T*) is the matrix $\left((-1)^{i+j} M_{ji}\right)$ and M_{ij} is the minor (determinant) of T_{ij}.

Theorem. (p.260 – no proof given) Matrix A is singular iff Det A = 0.

Proof. From [13.19], A is non-singular iff Det A ≠ 0. ■

Definition. $\mathcal{P}_{1\ldots n}$ is the **set of permutations of (1, …, *n*)**.

Theorem. [Bud] Det $T = \sum_{\pi\in\mathcal{P}_{12\cdots n}} \text{Sign}(\pi)\, T^1{}_{\pi(1)}\cdots T^n{}_{\pi(n)}$ (the standard definition)

Proof.

$$\text{Det } T = \frac{1}{n!}\varepsilon_{r\cdots s}\in^{t\cdots u} T^r{}_t\cdots T^s{}_u$$

$$= \frac{1}{n!}\sum_{\pi\in\mathcal{P}_{1\cdots n}}\sum_{\pi^*\in\mathcal{P}_{1\cdots n}}\varepsilon_{\pi^*(1)\cdots\pi^*(n)}\in^{\pi(1)\cdots\pi(n)} T^{\pi^*(1)}{}_{\pi(1)}\cdots T^{\pi^*(n)}{}_{\pi(n)}$$

(Replace Einstein notation.)

$$= \frac{1}{n!}\sum_{\pi\in\mathcal{P}_{1\cdots n}}\sum_{\pi^*\in\mathcal{P}_{1\cdots n}}\varepsilon_{\pi^*(1)\cdots\pi^*(n)}\in^{\pi(\pi^*(1))\cdots\pi(\pi^*(n))} T^{\pi^*(1)}{}_{\pi(\pi^*(1))}\cdots T^{\pi^*(n)}{}_{\pi(\pi^*(n))}$$

(Replace π by $\pi\circ\pi^*$ in $\in$ and *T*. The double sum over π and π^* is unchanged, in both expressions stepping over all permutations of (1, …, *n*), and the exponents of $\in$ continue to match the subscripts of *T*.)

$$= \frac{1}{n!}\sum_{\pi\in\mathcal{P}_{1\cdots n}}\sum_{\pi^*\in\mathcal{P}_{1\cdots n}}\mathbf{Sign}(\boldsymbol{\pi})\,\cancel{\varepsilon_{\pi^*(1)\cdots\pi^*(n)}\in^{\pi^*(1)\cdots\pi^*(n)}}\, T^{\pi^*(1)}{}_{\pi(\pi^*(1))}\cdots T^{\pi^*(n)}{}_{\pi(\pi^*(n))}$$

(Re-order superscripts of $\in$ by applying an inverse π permutation.)

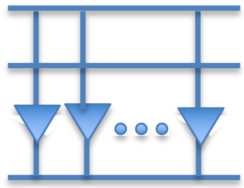
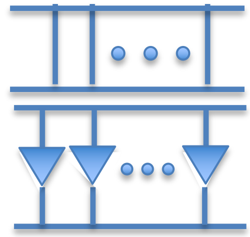
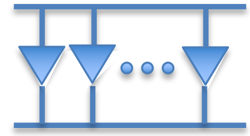
$$= \frac{1}{n!}\sum_{\pi\in\mathcal{P}_{1\cdots n}}\text{Sign}(\pi)\sum_{\pi^*\in\mathcal{P}_{1\cdots n}} T^{\mathbf{1}}{}_{\pi(\mathbf{1})}\cdots T^{\boldsymbol{n}}{}_{\pi(\boldsymbol{n})}$$

(This is just a simpler way to label the subscripts and superscripts of *T*. For example, if $\pi^*(3) = 1$ then $T^{\pi^*(3)}{}_{\pi(\pi^*(3))} = T^1{}_{\pi(1)}$.)

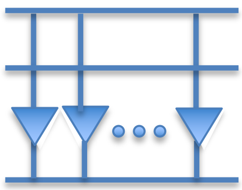
$$= \frac{n!}{n!}\sum_{\pi\in\mathcal{P}_{1\cdots n}}\text{Sign}(\pi)T^1{}_{\pi(1)}\cdots T^n{}_{\pi(n)}$$

$$= \sum_{\pi\in\mathcal{P}_{1\cdots n}}\text{Sign}(\pi)T^1{}_{\pi(1)}\cdots T^n{}_{\pi(n)} \quad ■$$

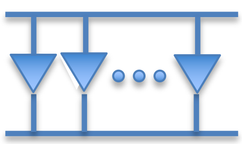
(See my solution to [13.21] for examples of this for *n* = 2 and 3.)

Theorem. [Bud]  =  = $n!$ 

Proof. Let $\mathcal{P}_{a\ldots g}$ be the set of permutations of (a, …, g). Then

 = $n!\varepsilon_{a\cdots g}\in^{r\cdots x}T^{[a}{}_{r}\cdots T^{g]}{}_{x}$

$$= \frac{n!}{n!}\varepsilon_{a\cdots g}\in^{r\cdots x}\sum_{\pi\in\mathcal{P}_{ab\cdots g}}\mathrm{Sign}(\pi)T^{\pi(a)}{}_{r}\cdots T^{\pi(g)}{}_{x}\overset{(*)}{=}n!\,\varepsilon_{a\cdots g}T^{a}{}_{r}\cdots T^{g}{}_{x}\in^{r\cdots x}$$

= $n!$ 

(*) π is the composition of transmutations (i.e., of pairwise permutations).

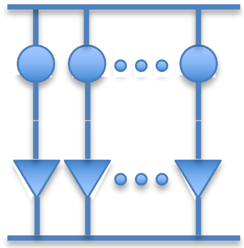
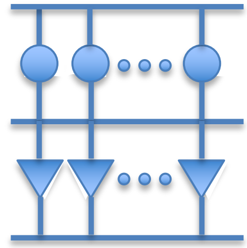
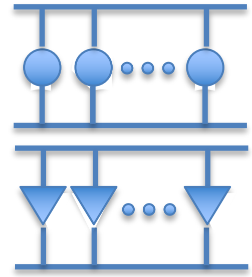
Let $\pi^*:\begin{cases} c\mapsto e \\ e\mapsto c\end{cases}$ be a transmutation. Then

$$\begin{aligned}
&\varepsilon_{a\cdots c\cdots e\cdots g}\in^{r\cdots t\cdots v\cdots x}\mathrm{Sign}(\pi^*)T^{\pi(a)}{}_{r}\cdots T^{\pi(c)}{}_{t}\cdots T^{\pi(e)}{}_{v}\cdots T^{\pi(g)}{}_{x}\\
&=\varepsilon_{a\cdots c\cdots e\cdots g}\in^{r\cdots t\cdots v\cdots x}\mathrm{Sign}(\pi^*)T^{a}{}_{r}\cdots T^{e}{}_{t}\cdots T^{c}{}_{v}\cdots T^{g}{}_{x}\\
&=\varepsilon_{a\cdots e\cdots c\cdots g}\in^{r\cdots t\cdots v\cdots x}\mathrm{Sign}(\pi^*)T^{a}{}_{r}\cdots T^{c}{}_{t}\cdots T^{e}{}_{v}\cdots T^{g}{}_{x}\ \text{(Rename } c\mapsto e\ \&\ e\mapsto c\text{)}\\
&=\mathrm{Sign}(\pi^*)\varepsilon_{a\cdots c\cdots e\cdots g}\in^{r\cdots t\cdots v\cdots x}\mathrm{Sign}(\pi^*)T^{a}{}_{r}\cdots T^{c}{}_{t}\cdots T^{e}{}_{v}\cdots T^{g}{}_{x}\\
&=\varepsilon_{a\cdots g}T^{a}{}_{r}\cdots T^{g}{}_{x}\in^{r\cdots x}.
\end{aligned}$$

So, for any permutation π, we have

$$\varepsilon_{a\cdots g}\in^{r\cdots x}\mathrm{Sign}(\pi)T^{\pi(a)}{}_{r}\cdots T^{\pi(g)}{}_{x}=\varepsilon_{a\cdots g}T^{a}{}_{r}\cdots T^{g}{}_{x}\in^{r\cdots x}\quad\blacksquare$$

Theorem. [13.22]

Det AB = $\frac{1}{n!}$  $= \left(\frac{1}{n!}\right)^2$  $= \left(\frac{1}{n!}\right)^2$ 

= Det A Det B

Corollary. T is nonsingular iff T is invertible, and Det T^{-1} = $\frac{1}{\text{Det } T}$.

Proof. $1 = \text{Det } I = \text{Det}\left[T\,T^{-1}\right] = \text{Det } T \text{ Det } T^{-1}$ ■

Corollary. If T is orthogonal, then Det T = ±1.

Proof. Clearly Det T^{T} = Det T. But $T^{T} \overset{\text{(defn)}}{=} T^{-1}$. So

$$\text{Det } T \overset{[13.22]}{=} \frac{1}{\text{Det }\left(T^{-1}\right)} = \frac{1}{\text{Det }\left(T^{\mathrm{T}}\right)} = \frac{1}{\text{Det } T}$$ ■

Definition. Vectors v and w are **orthogonal** if $v \bullet w = 0$. That is, the angle between them is 90°.

Theorem. A matrix is orthogonal (i.e., $T^{\mathrm{T}} = T^{-1}$) iff its column vectors are mutually orthogonal unit vectors.

Proof. Write the matrix T as $T = \left[t_1 \cdots t_n\right]$ where t_i are the column vectors. So

$$T^{\mathrm{T}} = \begin{bmatrix} t_1^{\mathrm{T}} \\ \vdots \\ t_n^{\mathrm{T}} \end{bmatrix}$$ where t_i^{T} are row vectors.

T is orthogonal

$$\Leftrightarrow \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} t_1^{\mathrm{T}} \\ \vdots \\ t_n^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} t_1 & \cdots & t_n \end{bmatrix} = \begin{bmatrix} t_1^{\mathrm{T}} t_1 & \cdots & t_1^{\mathrm{T}} t_n \\ \vdots & \ddots & \vdots \\ t_n^{\mathrm{T}} t_1 & \cdots & t_n^{\mathrm{T}} t_n \end{bmatrix}$$

$\Leftrightarrow$ $t_i^{\mathrm{T}} t_j = \delta_{ij}$; i.e., $t_i \perp t_j$ if $i \neq j$ and $\left\|t_i\right\|^2 = t_i^{\mathrm{T}} t_i = 1$

$\Leftrightarrow$ the column vectors are orthogonal unit vectors. ■

Example. $T = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}$ has orthogonal unit column vectors and

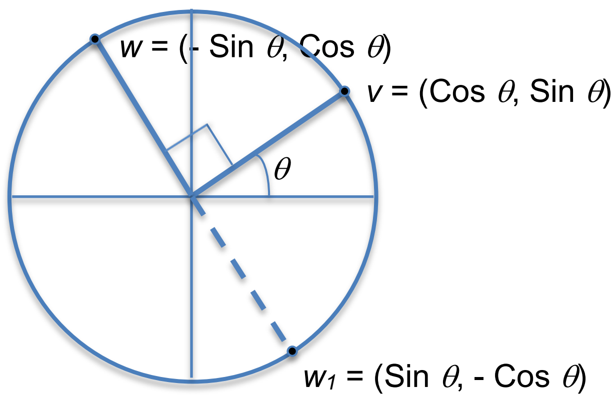
$$T\,T^{\mathrm{T}} = \frac{1}{10}\begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} = I\,.$$

Example. Orthogonal 2 x 2 Matrices: A and B

Let $A = \begin{pmatrix} v & w \end{pmatrix}$

$= \begin{pmatrix} \text{Cos}\,\theta & \text{- Sin}\,\theta \\ \text{Sin}\,\theta & \text{Cos}\,\theta \end{pmatrix}.$

$A^{\mathrm{T}} = \begin{pmatrix} \text{Cos}\,\theta & \text{Sin}\,\theta \\ \text{- Sin}\,\theta & \text{Cos}\,\theta \end{pmatrix}.$



$A^{\mathrm{T}} = A^{-1}$:

$$A\,A^{\mathrm{T}} = \begin{pmatrix} \text{Sin}^2\theta + \text{Cos}^2\theta & 0 \\ 0 & \text{Sin}^2\theta + \text{Cos}^2\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I} \quad ✔$$

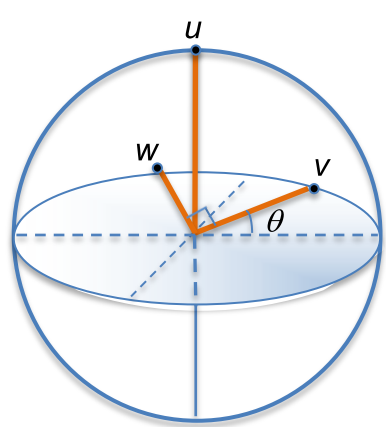
Similarly $A^{\mathrm{T}} A = \mathrm{I}$ ✔

So A is an orthogonal matrix ✔
Det A = Det A^{T} = $\text{Cos}^2\,\theta + \text{Sin}^2\,\theta = 1$ ✔
The column vectors of A are orthogonal: $v \perp w$ ✔

Let $B = \begin{pmatrix} \text{Cos}\,\theta & \text{Sin}\,\theta \\ \text{Sin}\,\theta & \text{- Cos}\,\theta \end{pmatrix} = B^{\mathrm{T}}$. Then $B\,B^{\mathrm{T}} = \mathrm{I}$, Det B = Det B^{T} = -1, and its column vectors v and w_1 are orthogonal.

Examples. Orthogonal 3 x 3 Matrices: A, B, and C

Let $v = \begin{pmatrix} \text{Cos}\,\theta \\ \text{Sin}\,\theta \\ 0 \end{pmatrix}$, $w = \begin{pmatrix} \text{- Sin}\,\theta \\ \text{Cos}\,\theta \\ 0 \end{pmatrix}$, and $u = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.



Let $\mathrm{A} = \begin{pmatrix} & & \\ v & w & u \\ & & \end{pmatrix} = \begin{pmatrix} \mathrm{Cos}\,\theta & -\mathrm{Sin}\,\theta & 0 \\ \mathrm{Sin}\,\theta & \mathrm{Cos}\,\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

$\mathrm{A}^{\mathrm{T}} = \begin{pmatrix} \mathrm{Cos}\,\theta & \mathrm{Sin}\,\theta & 0 \\ -\mathrm{Sin}\,\theta & \mathrm{Cos}\,\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$. A is orthogonal, its columns are orthogonal vectors, and its determinant is +1. ✔

Let $\mathrm{B} = \begin{pmatrix} \mathrm{Cos}\,\theta & -\mathrm{Sin}\,\theta & 0 \\ \mathrm{Sin}\,\theta & \mathrm{Cos}\,\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}$. B is orthogonal and its determinant is -1. ✔

Let C be the rotation of the orthogonal matrix A about the axis $\left\{t(a,b,c)\text{: } 0 < t < \infty,\ a^2+b^2+c^2=1\right\}$:

$$\mathrm{C} = \begin{pmatrix} \frac{1}{2}\left[1+a^2-b^2-c^2+\left(1-a^2+b^2+c^2\right)\mathrm{Cos}\,\theta\right] & 2\mathrm{Sin}\,\frac{\theta}{2}\left(-c\,\mathrm{Cos}\,\frac{\theta}{2}+ab\,\mathrm{Sin}\,\frac{\theta}{2}\right) & 2\mathrm{Sin}\,\frac{\theta}{2}\left(b\,\mathrm{Cos}\,\frac{\theta}{2}+ac\,\mathrm{Sin}\,\frac{\theta}{2}\right) \\ 2\mathrm{Sin}\,\frac{\theta}{2}\left(c\,\mathrm{Cos}\,\frac{\theta}{2}+ab\,\mathrm{Sin}\,\frac{\theta}{2}\right) & \frac{1}{2}\left[1-a^2+b^2-c^2+\left(1+a^2-b^2+c^2\right)\mathrm{Cos}\,\theta\right] & 2\mathrm{Sin}\,\frac{\theta}{2}\left(-a\,\mathrm{Cos}\,\frac{\theta}{2}+bc\,\mathrm{Sin}\,\frac{\theta}{2}\right) \\ 2\mathrm{Sin}\,\frac{\theta}{2}\left(-b\,\mathrm{Cos}\,\frac{\theta}{2}+ac\,\mathrm{Sin}\,\frac{\theta}{2}\right) & 2\mathrm{Sin}\,\frac{\theta}{2}\left(a\,\mathrm{Cos}\,\frac{\theta}{2}+bc\,\mathrm{Sin}\,\frac{\theta}{2}\right) & \frac{1}{2}\left[1-a^2-b^2+c^2+\left(1+a^2+b^2-c^2\right)\mathrm{Cos}\,\theta\right] \end{pmatrix}$$

It can be directly verified that C is an orthogonal matrix with mutually orthogonal column vectors and determinant +1. ✔

Definition. A **symmetry** of a vector space (V,+) is a non-singular linear transformation $T : \mathrm{V} \mapsto \mathrm{V}$. That is, T is 1-1, onto, and preserves the vector space structure: $T(a\,v + b\,w) = a\,Tv + b\,Tw$

Definition. The **General Linear Group GL(*n*)** is the group of symmetries of an *n*-dimensional vector space.

Theorem. GL(n) is the group of non-singular (n x n) matrices.

Proof. Let $T \in$ GL(n). Since T is a symmetry, T is non-singular. By [13.18] T is invertible. Thus, in any basis, T is represented by a non-singular (invertible) matrix. ■

Definition. The **Special Linear Group SL(*n*)** is the subset of GL(n) having determinant = 1.

Theorem. SL(n) is a normal subgroup of GL(n).

Proof. First, SL(n) is a group:

Closed: If $S_1, S_2 \in SL(n)$, then $\mathrm{Det}(S_1 S_2) = \mathrm{Det}(S_1)\,\mathrm{Det}(S_2) = 1$

$\Rightarrow S_1 S_2 \in SL(n)$.

Identity: $\mathrm{Det}(I) = 1 \Rightarrow I \in SL(n)$

Inverse: $1 = \mathrm{Det}(I) = \mathrm{Det}(S_1 S_1^{-1}) = \mathrm{Det}(S_1)\,\mathrm{Det}(S_1^{-1}) = \mathrm{Det}(S_1^{-1})$

$\Rightarrow S_1^{-1} \in SL(n)$

Also, SL(n) is normal:

Let $S \in SL(n)$ and $G \in GL(n)$. Then

$\mathrm{Det}(G^{-1} S G) = \mathrm{Det}(G^{-1})\,\mathrm{Det}(S)\,\mathrm{Det}(G) = \mathrm{Det}(G^{-1})\,\mathrm{Det}(G)$

$= \mathrm{Det}(G G^{-1}) = \mathrm{Det}(I) = 1$

$\Rightarrow G^{-1} S G \in SL(n) \Rightarrow G^{-1}\,SL(n)\,G = SL(n)$ ■

The groundwork has now been laid to introduce the table, below, that shows the relationships between SO(3), O(3), SL(3), GL(3), general linear transformations, orthogonality, determinants, and symmetries. The table shows that $SO(3) \subset O(3) \subset GL(3) \subset \mathcal{A}(\mathbb{R}^3)$ and $SO(3) \subset SL(3) \subset GL(3)$. It shows that GL(3) is both the set of symmetries of $\mathbb{R}^3$ and the set of non-singular matrices.

$\mathcal{A}(\mathbb{R}^3)$ = 3 x 3 Real Matrices

Vector Space of Linear Transformations on $\mathbb{R}^3$

Determinant	**Orthogonal**	**Unit sphere maps to a …**	**Matrix Type**
0	No	Circle, Ellipse, line segment or point	Singular
Between -1 and 0	No	Contracted reflected sphere or ellipsoid	**GL(3)** Non-singular Symmetries of $\mathbb{R}^3$
Between 0 and +1	No	Contracted sphere or ellipsoid	
-1	Yes	Reflected sphere	
	No	Reflected ellipsoid O(3)	
+1	Yes	SO(3) = sphere SL(3)	
	No	Ellipsoid	
< -1	No	Expanded reflected sphere or ellipsoid	
> 1	No	Expanded sphere or ellipsoid	

In general, non-singular matrices rotate and/or reflect and then squeeze and stretch the resultant unit sphere into an ellipsoid. However, singular matrices are more severe. They rotate and/or reflect and then squash the unit sphere down to a 2-dimensional circle or ellipse or even to a line or a point.

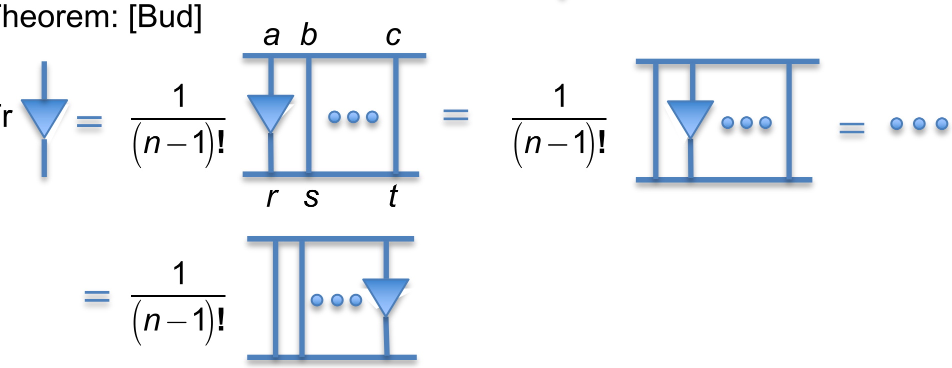
Orthogonal matrices preserve the sphere without squeezing or stretching any portion of it. This is achieved by limiting its operation to rotations. Anti-orthogonal matrices (det = -1) rotate and reflect. Matrices with orthogonal column vectors but determinant ≠ ±1 also expand or contract the sphere but in a uniform manner.

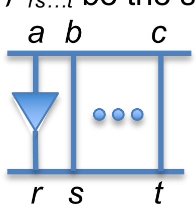
Consider a matrix that contains non-orthogonal vectors. In such a case the angle between the 1st and 2nd column vectors might be less than 90°, squeezing the sphere along associated plane. The angle between the 2nd and 3rd vectors would then be greater than 90°, stretching the sphere along that plane.

Matrices with positive determinant act on the sphere. Matrices with negative determinant behave exactly the same but act on the reflected sphere.

Definition. The **Trace** of A is T_r(A) = Tr  $= T^k{}_k = T^1{}_1 + \cdots T^n{}_n$.

Theorem: [Bud]

Tr 

Proof. Let $\mathcal{P}_{ab\ldots c}$ and $\mathcal{P}_{rs\ldots t}$ be the sets of permutations of (a,b,…,c) and (r,s,…,t), respectively. Let B = 

$$= \in^{rs\cdots t} \varepsilon_{ab\cdots c} T^a{}_r \delta^b_s \cdots \delta^c_t = \sum_{\pi\in\mathcal{P}_{ab\cdots c}} \sum_{\pi'\in\mathcal{P}_{rs\cdots t}} \in^{\pi'(r)\pi'(s)\cdots\pi'(t)} \varepsilon_{\pi(a)\pi(b)\cdots\pi(c)} T^{\pi(a)}{}_{\pi'(r)} \delta^{\pi(b)}_{\pi'(s)} \cdots \delta^{\pi(c)}_{\pi'(t)}.$$

Fix π. The only non-zero term in the sum is

$$\in^{\pi(a)\pi(b)\cdots\pi(c)} \varepsilon_{\pi(a)\pi(b)\cdots\pi(c)} T^{\pi(a)}{}_{\pi(a)} \delta^{\pi(b)}_{\pi(b)} \cdots \delta^{\pi(c)}_{\pi(c)} = T^{\pi(a)}{}_{\pi(a)}.$$

I showed in Problem [13.22] that $\in^{xy\cdots z} \varepsilon_{xy\cdots z} = 1$ for any fixed (x,y,…,z).

Thus, B = $\sum_{\pi \in \mathcal{P}_{ab\cdots c}} T^{\pi(a)}{}_{\pi(a)}$. This sum has n! terms composed of (n – 1)! terms equal to $T^a{}_a$, (n – 1)! terms equal to $T^b{}_b$, …, and (n – 1)! terms equal to $T^c{}_c$. So,

B = (n – 1)! ($T^a{}_a$ + $T^b{}_b$ + … + $T^c{}_c$) = (n – 1)! Tr (A) = (n – 1)! Tr 

Similarly for the other figures. ■

Theorem. [13.24] $\mathrm{Det}(\mathrm{I} + \in \mathrm{A}) = 1 + \in \mathrm{Tr}(\mathrm{A})$ if we ignore 2nd order and higher $\in$ terms.

Definition. $e^A \equiv \sum_{k=1}^{\infty} \frac{1}{k!} A^k$

Theorem. [13.25] Det e^A = $e^{\mathrm{Tr}(A)}$.

Definition. Let T be a linear transformation on a complex vector space V. An **eigenvector** is a non-zero vector v for which $\exists \lambda \in \mathbb{C}$ such that $Tv = \lambda v$ or, equivalently, $(T - \lambda I)v = 0$. λ is called an **eigenvalue.**

Note. $(T - \lambda I)v = 0$ and $v \neq 0$ $\Rightarrow$ $\mathrm{Det}(T - \lambda I) = 0$ $\Rightarrow$ $(T - \lambda I)$ is singular $\Rightarrow \lambda$ is a root of $\mathrm{Det}(T - \lambda I)$. That is, any root of $\mathrm{Det}(T - \lambda I)$ is an eigenvalue.

Definition. $p_T(\lambda) = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$ is called the **characteristic polynomial** of T. For some constant K, $p_T(\lambda) = K\,\mathrm{Det}(T - \lambda I)$.

Theorem. $p_T(\lambda) = 0$ is a polynomial equation of degree n. So T has n eigenvalues, possibly all the same. Thus, the linear transformation T has at least 1 eigenvector. Also, in general $\lambda \in \mathbb{C}$ even if T is a real matrix.

Definition. ***λ* has multiplicity *r*** means that λ appears r times in the characteristic polynomial. Eigenvalue multiplicities are called **degeneracies** in Quantum Mechanics.

Definition. The set of eigenvectors corresponding to λ generates a linear space called the **eigenspace of *λ*.**

Theorem. If d is the dimension of the eigenspace of λ and r is the multiplicity of λ then $1 \le d \le r$

Theorem. Let $\{\lambda_i\}$ be the set of eigenvalues of an n x n matrix T, and let r_i be the multiplicity of λ_i. Then $\sum r_i = n$.

Proof. Write the characteristic polynomial as $p_T(\lambda) = (-1)^n (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_m)^{r_m}$. Since there are n eigenvalues, $\sum r_i = n$. ■

Definition. Let T be an n x n *matrix and* S *any nonsingular* n x n matrix. Then T and $S^{-1}TS$ are called **similar matrices.**

Theorem. λ is an eigenvalue of T iff it is an eigenvalue of STS^{-1}.

Proof. Let x be an eigenvector of T corresponding to λ. Set $y = Sx \neq 0$. Then

$Tx = \lambda x \Leftrightarrow \lambda y = \lambda Sx = S\lambda x = STx = STS^{-1}Sx = STS^{-1}y$

$\Leftrightarrow$ y is an eigenvector of STS^{-1} corresponding to λ

$\Leftrightarrow$ λ is an eigenvalue of STS^{-1}. ■

Theorem. [13.30] Suppose $\{e_k\}$ and $\{f_k\}$ are bases for a vector space V, and $f_k = Te_k$. Then

$$f_j = \begin{pmatrix} T^1{}_j \\ \vdots \\ T^n{}_j \end{pmatrix}.$$

That is, the components of f_j in basis $\{e_k\}$ are $(T^1{}_j, \cdots, T^n{}_j)$.

Theorem. [13.31] If the eigenspace dimension of every multiple eigenvalue equals its multiplicity, then there is a basis for V composed of eigenvectors, and the matrix of T in this basis is

$$T = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

The next theorem states that even when the hypothesis of the above theorem is not satisfied, the matrix of T can at least be written in upper triangular form.

Theorem. (Note 13.12): **Jordan Canonical Form:** Let $\{\lambda_i\}$ be the set of eigenvalues of an *n* x *n* matrix *T*, and let r_i be the multiplicity of λ_i. Then there is a basis for V such that the matrix of *T* in this basis is as shown at the right.

$$\left(\begin{array}{c|c|c}
\begin{matrix}\lambda_1 & 1 & & & \\ & \lambda_1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_1\end{matrix} & 0 & \begin{matrix}\cdots & 0 \\ \iddots & \vdots\end{matrix} \\ \hline
 & \begin{matrix}\lambda_2 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_{n-1}\end{matrix} & 0 \\ \hline
\begin{matrix}\vdots & \iddots \\ 0 & \cdots\end{matrix} & & \begin{matrix}\lambda_n & 1 & & \\ & \lambda_n & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_n\end{matrix}
\end{array}\right)$$

Definition. An $n \times n$ matrix *T* is Hermitian if $t_{ji} = \overline{t_{ij}}$ (complex conjugate).

Theorem. Let *T* be an $n \times n$ Hermitian matrix. Then the eigenspace dimension of every multiple eigenvalue equals its multiplicity. Thus there is a basis for *T* composed of eigenvectors and the matrix of *T* in this basis is diagonal.

Representations and Lie Algebras

Definition. Let $T : \mathrm{G} \to \mathcal{G}$ be a homomorphism of a group G to some well-known standard group $\mathcal{G}$. The image *T*(G) is called a **Group Representation of G**. In this section we take $\mathcal{G}$ to be GL(*n*), the multiplicative group of non-singular *n* × *n* matrices. *T* is **faithful** if it is 1-1.

Theorem. [13.32] Every finite group has a faithful representation. Every finite dimensional Lie group has a faithful representation.

GL(*n*) is not Abelian. Since finite groups include both Abelian and non-Abelian, this means that GL(*n*) has some Abelian subgroups.

Example. Let G be the additive group of modulus 3: G = {0, 1, 2}. Then *T*(G) is a group representation where

$$T(0) = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} = I, \quad T(1) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad T(2) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Since $T(1)T(2) = I = T(2)T(1)$, $\{T(0), T(1), T(2)\}$ is an Abelian subgroup of GL(3).

We will see in Chapter 14 that the theory of representations of continuous groups by linear transformations can be converted to the study of representations of Lie algebras, which we define next.

Definition. An **algebra** is ring R that is also a vector space (that is, it has scalar multiplication in addition to addition and regular multiplication) and that for all a, b in R and scalar α we have $\alpha(a\,b) = (\alpha a)b = a(\alpha b)$. If the underlying ring is associative, then it is an **associative algebra**.

Example. If V is a vector space then the set of linear transformations, $\mathcal{A}$(V), is an algebra. E.g., $\mathcal{A}(\mathbb{R}^3)$, the set of 3 × 3 matrices, is an algebra: you can add and multiply matrices as well as multiply them by scalars.

Example. GL(3) is not an algebra nor even a ring because it is not closed under matrix addition. It is just a multiplicative group. For example, addition of 2 non-singular matrices can yield the zero matrix that is singular and not in GL(3).

I next give the standard definition of a Lie algebra. I will shortly prove that Penrose's definition (of a special case) satisfies this Lie algebra definition.

Definition. (Standard definition) ($\mathfrak{g},+,[\bullet,\bullet]$) is a **Lie algebra** if $(\mathfrak{g},+)$ is a vector space over a field F and a **Lie bracket** binary operator exists and satisfies

- **Bilinearity:**
  $[ax+by,z]=a[x,z]+b[y,z]$ and $[z,ax+by]=a[z,x]+b[z,y]$
  $\forall a,b\in F$ and $x,y,z\in\mathfrak{g}$
- **Alternativity:** $[x,x]=0\ \forall x\in\mathfrak{g}$
- **Jacobi Identity:** $[x,[y,z]]+[z,[x,y]]+[y,[z,x]]=0\ \forall x,y,z\in\mathfrak{g}$

Note that bilinearity and alternativity imply

- **Anticommutativity**: $[x,y] = -[y,x]$:
  $$0=[x+y,\,x+y]=\cancel{[x,x]}+[x,y]+[y,x]+\cancel{[y,y]} \qquad ✔$$

Theorem. A Lie algebra is an algebra.

Proof. We have only to show that $(\mathfrak{g},+,[\bullet,\bullet])$ is a ring, and only the bracket distributive property remains to be shown. But left and right bracket distributivity follow from bilinearity. ■

Example. $(R^3, +, \times)$ is a Lie algebra where × is the cross product.

Theorem. [13.34] Let A, B $\in \mathfrak{g}$. Then

(a) $(I+\in A)(I+\in B)=I+\in(A+B)$ if we ignore terms $o(\in)^2$

(b) $(I+\in A)(I+\in B)(I+\in A)^{-1}(I+\in B)^{-1}=I+\in^2[A,B]$ if we ignore $o(\in)^3$

If "infinitesimals" *a* and *b* are represented by $(I+\in A)$ and $(I+\in B)$, then we see that the product $aba^{-1}b^{-1}$ is represented by Lie brackets. Thus we make the following definition.

Definition. Let G be a group. $\{aba^{-1}b^{-1} : a,b \in G\}$ is the **set of group commutators.**

Definition. (Penrose's definition) Let $\mathcal{G}$ be a subgroup of GL(*n*). Let $(\mathcal{G}^{*},\bullet)$ be the vector space generated from $\mathcal{G}$ by the addition of scalar multiplication. For A, B $\in \mathcal{G}^*$, a **Lie bracket** is **[A,B] ≡ AB – BA** (commutator operation). Note that [A,B] is not necessarily in $\mathcal{G}$ because $\mathcal{G}$ is not required to be closed under subtraction. Let $\mathcal{G}^{**}$ be the set of commutators generated from $\mathcal{G}^*$; i.e., $\mathcal{G}^{**}=\{[A,B] : A,B \in \mathcal{G}^*\}$. The **Lie algebra generated by $\mathcal{G}$** is the algebra $\mathfrak{g}$ generated from $\mathcal{G}^{**}$ by applying +, –, and Lie bracket operations repeatedly until nothing new occurs. Again, note that $\mathcal{G}$ is not necessarily contained in $\mathfrak{g}$.

Construction. We generate $\mathfrak{g}$ inductively:

$$\mathfrak{g}_0=\mathcal{G}^{**} \qquad \text{Level 0}$$

$$\mathfrak{g}_1=\{A\pm B : A,B\in \mathfrak{g}_0\}\cup\{[A,B] : A,B\in \mathfrak{g}_0\} \qquad \text{Level 1}$$

$$\mathfrak{g}_2=\{A\pm B : A,B\in \mathfrak{g}_1\}\cup\{[A,B] : A,B\in \mathfrak{g}_1\} \qquad \text{Level 2}$$

$$\vdots$$

$$\mathfrak{g}_\infty=\bigcup_{k=1}^{\infty}\mathfrak{g}_k \qquad \text{Level } \infty$$

Claim: $\mathfrak{g}=\mathfrak{g}_\infty$:

Let *A*, *B* $\in \mathfrak{g}_\infty$. Then $\exists$ *n* such that *A*, *B* $\in \mathfrak{g}_n$. Thus *A* ± *B* $\in \mathfrak{g}_{n+1}\subset \mathfrak{g}_\infty$ and $[A,B]\in \mathfrak{g}_{n+1}\subset \mathfrak{g}_\infty$. So $\mathfrak{g}_\infty$ is closed under +, −, and $[\bullet,\bullet]$. ✔

Example. Let $\mathcal{G}$ = SL($n,\mathbb{R}$), the multiplicative group of real matrices having determinant 1. Show $\mathfrak{g}$ is the set of *n* × *n* matrices having trace 0.

Solution.

$\mathcal{G} = \{A\text{: det } A = 1\}$.

$\mathcal{G}^* = \left\{\alpha A : \alpha \in \mathbb{R} \text{ and } \det A = 1\right\}$

$$\mathfrak{g}_0 = \mathcal{G}^{**} = \left\{ \left[C,D\right] : C, D \in \mathcal{G}^*\right\} = \left\{ \left[\alpha A, \beta B\right] : \alpha, \beta \in \mathbb{R} \text{ and } A,B \in \mathcal{G}\right\}$$
$$= \left\{\alpha\beta\left[A,B\right] : \alpha, \beta \in \mathbb{R} \text{ and } A,B \in \mathcal{G}\right\} = \left\{\alpha\left[A,B\right] : \alpha \in \mathbb{R} \text{ and } A,B \in \mathcal{G}\right\}.$$

First, observe that if $A = \left(a_{ij}\right)$ and $B = \left(b_{ij}\right)$ are any matrices then

$$AB = \left(\sum_k a_{ik} b_{kj}\right), \; BA = \left(\sum_k a_{kj} b_{ik}\right)$$
$$\Rightarrow \quad \text{Tr } AB = \sum_{i,k} a_{ik} b_{ki}$$

and

$$\text{Tr } BA = \sum_{i,k} a_{ki} b_{ik} = \sum_{i,k} a_{ik} b_{ki} = \text{Tr } AB. \quad ✔$$

Therefore for all A, B :

$$\text{Tr } \left[A,B\right] = \text{Tr } \left(AB - BA\right) = \text{Tr } AB - \text{Tr } BA = 0. \quad ✔$$

So, let $D = \left[A, B\right] \in \mathfrak{g}_0 = \mathcal{G}^{**} \Rightarrow \text{Tr}\left(D\right) = 0$. We proceed by induction. Suppose $D \in \mathfrak{g}_k \Rightarrow \text{Tr}\left(D\right) = 0$ and let $A \in \mathfrak{g}_{k+1}$. Either $A = B \pm C$ or $A = [B,C]$ where $B,C \in \mathfrak{g}_k$. By induction hypothesis $\text{Tr}\left(B\right) = 0 = \text{Tr}\left(C\right)$. If $A = B \pm C$ then $\text{Tr}\left(A\right) = \text{Tr}\left(B\right) \pm \text{Tr}\left(C\right) = 0$. If $A = [B,C]$ then $\text{Tr}\left(A\right) = \text{Tr } \left[B,C\right] = 0$. Therefore $\text{Tr}\left(A\right) = 0 \quad \forall A \in \mathfrak{g}$. ✔

Conversely, suppose Tr $A = 0$. $A \in \mathfrak{g}$ if for some n, $A \in \mathfrak{g}_{n+1}$. To show that $A \in \mathfrak{g}_{n+1}$ we wish to find matrices $B,C \in \mathfrak{g}_n$ such that either $A = B \pm C$ or $A = [B,C]$. That seems hard to do. It is not difficult to show for $\mathcal{G} = \text{SL}(n,\mathbb{R})$ that $\mathfrak{g} = \left\{A : \exists p \in \mathbb{Z}^+ \text{ and } A_k \in \mathfrak{g}_0 \text{ such that } A = A_1 + \cdots + A_p\right\}$. It might be easier to find $A_k \in \mathfrak{g}_0$ such that $A = A_1 + \cdots + A_n$. By definition, $A_k = \left[\beta_k B_k, \gamma_k C_k\right]$ where det B_k = det C_k = 1. However, I cannot figure out how to find $A_1 + \ldots + A_n$. ■

Lemma. [13.35] Let A, B $\in \mathfrak{g}$ and $\lambda \in \mathbb{C}$. Then

(a) [A+B,C] = [A,C] + [B,C] and [λA,B] = λ [A,B] (Lie bracket left distributivity)

(b) [B,A] = - [A,B] (Lie bracket antisymmetry, also called anticommutativity)

(c) [A, [B,C]] + [B, [C,A]] + [C, [A,B]] = 0 (Jacobi identity)

(d) dim $\mathcal{G}^*$ ≤ dim G . If T is faithful, then dim $\mathcal{G}^*$ = dim G

Theorem. $\left(\mathfrak{g},+,[\bullet,\bullet]\right)$ is an ~~associative~~ Lie algebra.

Proof. We must show that $\mathfrak{g}$ Is a vector space, that bracket satisfies bilinearity, alternativity, and the Jacobi identity, and that the algebra is associative.

($\mathfrak{g}$, +) is an Abelian group:

- We just showed that $\mathfrak{g}_\infty$ is closed under +
- [0] $\in \mathfrak{g}$ because for any $A \in \mathfrak{g}$ we have [0] = $A - A$
- - A $\in \mathfrak{g}$ because $\mathfrak{g}_\infty$ is closed under subtraction
- $\mathfrak{g}$ is commutative because + operates on matrices

($\mathfrak{g}$, +) is a vector space over $\mathbb{R}$ (or $\mathbb{C}$):

First we must confirm that if α is a scalar and A $\in \mathfrak{g}$ then αA$\in \mathfrak{g}$. This is true for Level 0 since $\mathfrak{g}^*$ is a vector space and $\mathfrak{g}^{**}$ is constructed from $\mathfrak{g}^*$. Using simple induction, the same argument shows that if it is true for level k then it is true for level k+1.

Next, let α and β be scalars and A, B, $C \in \mathfrak{g}$.

(1) $\alpha(A+B)=\alpha A+\alpha B$ is true because A and B are matrices.

Similarly,

(2) $(\alpha+\beta)A=\alpha A+\beta A$,

(3) $\alpha(\beta A)=(\alpha\beta)A$, and

(4) $1A=A$.

Bilinearity:

$$[\alpha A+\beta B, C]=(\alpha A+\beta B)C-C(\alpha A+\beta B)=\alpha(AC-CA)+\beta(BC-CB)$$
$$=\alpha[A,C]+\beta[B,C]$$

because A, B, C are matrices. Similarly right bilinearity holds

Alternativity:

[A, A] = A A – A A = 0

Jacobi Identity: (part of lemma, but here is the proof)

$$[A,[B,C]]+[B,[C,A]]+[C,[A,B]]$$
$$=[A(BC-CB)-(BC-CB)A]+[B(CA-AC)-(CA-AC)B]$$
$$+[C(AB-BA)-(AB-BA)C]$$
$$=[ABC-ACB-BCA+CBA]+[BCA-BAC-CAB+ACB]$$
$$+[CAB-CBA-ABC+BAC]=0$$

■

~~The ring $(\mathfrak{g},+,[\bullet,\bullet])$ is associative~~:

We must show $[A,[B,C]]=[[A,B],C]$ for $A,B,C\in\mathfrak{g}$.

$$[A,[B,C]]=A(BC-CB)-(BC-CB)A=ABC-ACB-BCA+CBA$$

$$[[A,B],C]=(AB-BA)C-C(AB-BA)=ABC-BAC-CAB+CBA \quad ■$$

Convention. Henceforth we assume T is faithful. Thus, dim $\mathcal{G}^*$ = n = dim $\mathcal{G}$.

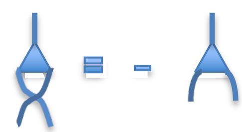
Definition. Let n be the dimension of the vector space $\mathcal{G}^*$ and $(E_1, E_2, \ldots, E_n)$ a basis for $\mathcal{G}^*$. Then

$$\exists\gamma_{\alpha\beta}{}^{\chi} \text{ where } \alpha,\beta,\chi\in\{1,\ 2,\ \cdots,\ n\} \text{ such that } [E_\alpha,E_\beta]=\gamma_{\alpha\beta}{}^{\chi}E_\chi .$$

The n^3 components $\gamma_{\alpha\beta}{}^{\chi}$ are called the **structural constants for G** and can be expressed in diagrammatic form as shown at right.

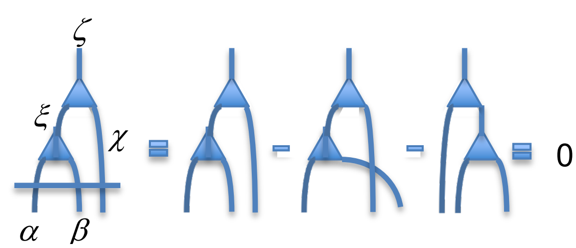
The $\gamma_{\alpha\beta}{}^{\chi}$ are not all independent because they satisfy relations in the next theorem.

Theorem. [13.36] $\gamma_{\beta\alpha}{}^{\chi}=-\gamma_{\alpha\beta}{}^{\chi}$ and $\gamma_{[\alpha\beta}{}^{\xi}\,\gamma_{\chi]\xi}{}^{\zeta}=0$



Proof. This follows from Lie bracket antisymmetry and the Jacobi identity.
■

This theorem can be expressed in diagrammatic form as shown.



Definition. Let V be a vector space.
The **dual space V*** is defined to be $V^*=\{f:V\to\mathbb{R} \text{ or } \mathbb{C}: f \text{ is a linear map}\}$. By convention the vector space consists of column vectors $x=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$. The next theorem says that the dual space consists of row vectors $y=[y_1\cdots y_n]$.

Theorem. Let V be a vector space (of column vectors) and $V^T = [y : y \text{ is a row vector and } y^T \in V]$ the corresponding vector space of row vectors. (Recall that the superscript T means transpose.) Then

(1) $V^* = \{f_y : V \to \mathbb{R} \text{ or } \mathbb{C} :\ f_y : x \mapsto y\,x \text{ for } x \in V \text{ and } y \in V^T\}$

and

(2) $V^* \cong V^T$.

Proof. Let $x \in V$ denote a column vector: $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. Consider the basis for V composed of $\{e_1, e_2, \ldots, e_n\}$ *where* $e_k = \begin{bmatrix} 0 \\ \vdots \\ 1_k \\ \vdots \\ 0 \end{bmatrix}$.

Let $f \in V^*$. Define $y_k = f(e_k)$ for all k and set $y = [y_1, \cdots, y_n]$. Then

$$f(x) = f\left(\sum_{k=1}^{n} x_k\, e_k\right) = \sum_{k=1}^{n} f(x_k\, e_k) = \sum_{k=1}^{n} x_k\, f(\,e_k) = \sum_{k=1}^{n} x_k\, y_k$$

$$= [y_1, \cdots, y_n]\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = y\,x$$

$$= f_y(x)$$

Thus every linear map $f : V \to \mathbb{R}$ or $\mathbb{C}$ equals f_y for some $y \in V$, namely $y = (\,f_y(e_1), \cdots, f_y(e_n)\,)$. Hence, $V^* \subseteq \{f_y : y \in V^T\}$. Also, every map f_y is linear:

$$f_y(\alpha v_1 + \beta v_2) = y(\alpha v_1 + \beta v_2) = \alpha y v_1 + \beta y v_2 = \alpha f_y(v_1) + \beta f_y(v_2) \quad ✔$$

That is, $\{f_y : y \in V^T\} \subseteq V^*$. So, $V^* = \{f_y : y \in V^T\}$, proving (1) ✔

The mapping $f : V^T \to V^* : f(y) = f_y$ is an isomorphism because $\forall\, x \in V$

$$[f(y+z)](x) = f_{y+z}(x) = (y+z)x = y\,x + z\,x = f_y(x) + f_z(x) = [f(y) + f(z)](x)$$

for y, $z \in V^T$, and for all scalars α

$$[f(\alpha y)](x) = f_{\alpha y}(x) = \alpha\, y\, x = \alpha f_y(x) = [\alpha f(y)](x) \quad ■$$

Note: There are duality theories for groups and rings. In these theories the first dual is a function space and the second dual turns out to be isomorphic to the original structure. But with vector spaces Penrose has proved that $V^* \cong V^{\mathsf{T}}$, and clearly $V^{\mathsf{T}} \cong V$, resulting in $V^* \cong V$, the first dual isomorphic to the original vector space. This occurs because the product of a row vector and a column vector is a scalar, as required. With groups and rings, no operation can result in a scalar, an element that is outside the original structure. So the first dual cannot be isomorphic to the original for groups and rings.

Definition. Let G be a group. A vector space V is called a **representation space for G** if G is represented by a group $\mathcal{G}$ of linear transformations on V (i.e., a group of invertible matrices that act on vectors in V).

Let $T \in \mathcal{G}$. For x $\in$ V, $T : \mathrm{V} \to \mathrm{V}$ can be written as $x \mapsto Tx$, or in matrix form as $x^a \mapsto T^a{}_b x^b$ where $T^a{}_b$ is an n x n matrix and x^a and x^b are column n-vectors.
Set

$$S = T^{-1}, \text{ or } S^a{}_b = \left(T^a{}_b\right)^{-1}.$$

Then

$$ST = I, \text{ or } S^a{}_b T^b{}_c = \delta^a_c .$$

Let $y \in$ V*. Then

$$y : \mathrm{V} \to \mathbb{R} \text{ or } \mathbb{C} : x \mapsto y\,x \text{ for } x \in \mathrm{V}.$$

The identity map

$$y\,x \mapsto y\,x = y(ST)x = (yS)(Tx)$$

can be decomposed into maps

$$y \mapsto yS, \text{ or } y_a \mapsto y_b S^b{}_a$$

and

$$x \mapsto Tx .$$

$$\begin{tikzcd} \mathrm{V} \arrow[r, "S"] \arrow[dr, dashed, "yS"'] & \mathrm{V} \arrow[d, "y"] \\ & \mathbb{R} \text{ or } \mathbb{C} \end{tikzcd}$$

The figure illustrates that both $y \in \mathrm{V}^*$ and $yS \in \mathrm{V}^*$. Also notice that x is a column vector and so is written to the right of the matrix T, and y is a row vector and is written to the left of the matrix S. We prefer to use column vectors so we often write

$$y^{\mathsf{T}} \mapsto S^{\mathsf{T}} y^{\mathsf{T}}, \text{ or } y^a \mapsto \left(S^{\mathsf{T}}\right)^a{}_b\, y^b$$

The mapping $y^{\mathsf{T}} \mapsto S^{\mathsf{T}} y^{\mathsf{T}}$ plays the central role in the representation-space theorem for tensors (in a few pages). In the representation-space theorem for the dual space, next, the mapping $(y \mapsto yS) : \mathrm{V}^* \to \mathrm{V}^*$ plays the central role.

Theorem. If V is a representation space for a group G, then so is V*.

Proof. The short proof is that this theorem is true because $V^* \cong V$. However, it is insightful to do a constructive proof.

By definition of representation, there is a subgroup $\mathcal{G} \subset \mathrm{GL}(n)$ and an isomorphism $T : G \to \mathcal{G}$ such that $\forall g \in G \;\; T(g) : \mathrm{V} \to \mathrm{V}$ is a linear transformation on V. We seek another subgroup $\mathcal{G}^* \subset \mathrm{GL}(n)$ and an isomorphism $\mathcal{T} : G \to \mathcal{G}^*$ such that $\forall g \in G \;\; \mathcal{T}(g) : \mathrm{V}^* \to \mathrm{V}^*$ is a linear transformation on V*.

Define

(1) $T_g = T(g)$ for $g \in$ G.

T_g is a linear transformation on V. Since T is an isomorphism defined on G,

(2) $T_{g_1 g_2} \overset{(1)}{=} T(g_1 g_2) = T(g_1)\, T(g_2) \overset{(1)}{=} T_{g_1} T_{g_2}$ for $g_1, g_2 \in G$.

Let

(3) $S_g = T_g^{-1}$.

S_g is also a linear transformation on V. Define a mapping

(4) $\mathcal{T}_g : \mathrm{V}^* \to \mathrm{V}^* : \mathcal{T}_g(y) = y S_g$ for $y \in \mathrm{V}^*$.

$$\begin{tikzcd} \mathrm{V} \arrow[r, "S_g"] \arrow[dr, dashed, "y S_g"'] & \mathrm{V} \arrow[d, "y"] \\ & \mathbb{R} \text{ or } \mathbb{C} \end{tikzcd}$$

Recall that y is a row vector as is yS. Define

$$\mathcal{G}^* = \{\mathcal{T}_g : g \in G\}$$

and

(5) $\mathcal{T} : G \to \mathcal{G}^* : \mathcal{T}(g) = \mathcal{T}_g$ for $g \in$ G.

$\mathcal{G}^*$ Is clearly a subgroup of GL(n). We need to show that $\mathcal{T}$ is an isomorphism. Let $g_1, g_2 \in$ G and set

(6) $g_3 = g_1 g_2$.

Then

(7) $\mathcal{T}_{g_3} \overset{(6)}{=} \mathcal{T}_{g_1 g_2} \overset{(2)}{=} \mathcal{T}_{g_1} \mathcal{T}_{g_2}$

which implies

(8) $\quad S_{g_3} \overset{(3)}{=} T_{g_3}{}^{-1} \overset{(7)}{=} T_{g_2}{}^{-1} T_{g_1}{}^{-1} \overset{(3)}{=} S_{g_2} S_{g_1}$.

Let $y \in$ V*. Then

$$\mathcal{T}(g_1 g_2)(y) \overset{(6)}{=} \mathcal{T}(g_3)(y) \overset{(5)}{=} \mathcal{T}_{g_3}(y) \overset{(4)}{=} y S_{g_3} \overset{(8)}{=} y S_{g_2} S_{g_1} \overset{(4)}{=} \left[\mathcal{T}_{g_2}(y)\right] S_{g_1}$$
$$\overset{(4)}{=} \mathcal{T}_{g_1}\left[\mathcal{T}_{g_2}(y)\right] \overset{(5)}{=} \mathcal{T}(g_1)\left[\mathcal{T}(g_2)(y)\right] = \left[\mathcal{T}(g_1)\mathcal{T}(g_2)\right](y),$$

or

$$\mathcal{T}(g_1 g_2) = \mathcal{T}(g_1)\mathcal{T}(g_2).$$

That is, $\mathcal{T}$ is a homomorphism ✔

To show that $\mathcal{T}$ is an isomorphism, we must show that it is 1-1.

$$\mathcal{T}(g_1) = \mathcal{T}(g_2) \quad \overset{(5)}{\Leftrightarrow} \quad \mathcal{T}_{g_1} = \mathcal{T}_{g_2} \quad \overset{(4)}{\Leftrightarrow} \quad y\, S_{g_1} = y\, S_{g_2} \quad \forall\, \text{y} \in \text{V*}$$
$$\Leftrightarrow \quad S_{g_1} = S_{g_2} \quad \overset{(3)}{\Leftrightarrow} \quad T_{g_1} = T_{g_2} \quad \overset{(1)}{\Leftrightarrow} \quad T(g_1) = T(g_2)$$
$$\Rightarrow \quad g_1 = g_2 \text{ since } T \text{ is an isomorphism.}$$

So $\mathcal{T}$ is 1-1. ✔ ■

Tensors

Definitions. Let V be an m-dimensional vector space over a field F_V and W be an n-dimensional vector space over a field F_W. Let $T : V \times W \to X$ be a map where X can be $V \times W$, $\mathbb{R}$, $\mathbb{C}$ or any algebraic space or set.

- ***T* is an outer product** if $T(v,w) = \sum_{i,j} v_i * w_j$ for some operation $*$
- ***T* is an inner product** if $m = n$ and $T(v,w) = \sum_{i=1}^{n} v_i * w_i$ for some operation $*$
- ***T* is bilinear** if V and W share a common field F (common scalars) and it is linear in V and W independently; i.e., T satisfies
  (1) $\alpha T(v,w) = T(\alpha v,w) = T(v,\alpha w)$
  (2) $T(u+v,w) = T(u,w) + T(v,w)$
  (3) $T(v,w+x) = T(v,w) + T(v,x)$

Conventions.

- Superscripts are used for column vectors and subscripts are used for row vectors. Technically, tensors are basis independent. So **superscript tensors** are called **vectors** and **subscript tensors** are called **covectors.**
- Einstein Summation Convention is used with tensors. That is, when an index appears in an expression as both a subscript and a superscript, then it represents a summation.

Example 1. $v_i w^i$ means $\sum_i v_i w^i$.

Example 2. $x_a y^b Q_b^a$ means $\sum_a \sum_b x_a y^b Q_b^a$.

Definition. The **tensor product of vectors** $v = (v^i)$ and $w = (w^j)$ is the bilinear outer product $\otimes$**:** $\mathrm{V} \times \mathrm{W} \to \mathrm{V} \otimes \mathrm{W}$:

$$\boldsymbol{v} \otimes \boldsymbol{w} = \begin{pmatrix} v^1 \\ \vdots \\ v^m \end{pmatrix} \otimes \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix} \equiv \begin{pmatrix} v^1 \otimes w^1 & \cdots & v^1 \otimes w^n \\ \vdots & & \vdots \\ v^m \otimes w^1{}_1 & \cdots & v^m \otimes w^n \end{pmatrix}$$

or

$$\boldsymbol{v} \otimes \boldsymbol{w} = \begin{pmatrix} v^1 w^1 & \cdots & v^1 w^n \\ \vdots & & \vdots \\ v^m w^1{}_1 & \cdots & v^m w^n \end{pmatrix} \text{ for short.}$$

Bilinear Product Rules: The following rules make $\otimes$ bilinear:

(1) $\alpha(v \otimes w) = (\alpha v) \otimes w = v \otimes (\alpha w)$, where α is a scalar.

(2) $(x + v) \otimes w = x \otimes w + v \otimes w$

(3) $v \otimes (y + w) = v \otimes y + v \otimes w$.

Tensor components may not be combined.

The components $v^i \otimes w^j = v^i w^j$ may not be multiplied out, interchanged, or combined in any way because they belong to different vector spaces. Also, in Quantum Mechanics, *v* is the state of one system and *w* the state of another system. So, combining them doesn't make sense.

Note 1. **Bilinear** means linear in each of V and W separately, a reminder that there is no mixing of V and W.

Note 2. **Multilinear** means linear in each of several vector spaces separately.

Note 3. Rules (1) – (3) can also be expressed in matrix form:

$$\textbf{(1)}\quad \alpha\begin{bmatrix} v^1w^1 & \cdots & v^1w^n \\ \vdots & & \vdots \\ v^mw^1_1 & \cdots & v^mw^n \end{bmatrix} = \begin{bmatrix} \left(\alpha v^1\right)w^1 & \cdots & \left(\alpha v^1\right)w^n \\ \vdots & & \vdots \\ \left(\alpha v^m\right)w^1 & \cdots & \left(\alpha v^m\right)w^n \end{bmatrix}$$

$$= \begin{bmatrix} v^1\left(\alpha w^1\right) & \cdots & v^1\left(\alpha w^n\right) \\ \vdots & & \vdots \\ v^m\left(\alpha w^1\right) & \cdots & v^m\left(\alpha w^n\right) \end{bmatrix}$$

$$\textbf{(2)}\quad \begin{bmatrix} \left(x^1+v^1\right)w^1 & \cdots & \left(x^1+v^1\right)w^n \\ \vdots & & \vdots \\ \left(x^m+v^m\right)w^1 & \cdots & \left(x^m+v^m\right)w^n \end{bmatrix}$$

$$= \begin{bmatrix} x^1w^1 & \cdots & x^1w^n \\ \vdots & & \vdots \\ x^mw^1 & \cdots & x^mw^n \end{bmatrix} + \begin{bmatrix} v^1w^1 & \cdots & v^1w^n \\ \vdots & & \vdots \\ v^mw^1 & \cdots & v^mw^n \end{bmatrix}$$

and similarly for (3).

Convention. Henceforth we assume that all vector spaces, V, W, etc., have the same field of scalars.

Definition. Let $\left\{e_i\right\}_{i=1}^m$ and $\left\{f_j\right\}_{j=1}^n$ be bases for V and W, respectively, where

$$\mathbf{e}_i = \begin{pmatrix} \mathbf{0} \\ \vdots \\ \mathbf{1}_i \\ \vdots \\ \mathbf{0} \end{pmatrix} \text{ and } \boldsymbol{f}_j = \begin{pmatrix} \mathbf{0} \\ \vdots \\ \mathbf{1}_j \\ \vdots \\ \mathbf{0} \end{pmatrix} \text{ are vectors. Then}$$

$$e_i \otimes f_j = \begin{array}{c} \\ 1 \\ \\ i \\ \\ m \end{array} \!\!\begin{array}{c} \begin{array}{ccccc} 1 & & j & & n \end{array} \\ \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \end{array}$$

is a **tensor basis element**.

We can express any vector tensor product in terms of the tensor basis:

$$v \otimes w = \begin{pmatrix} v^1 w^1 & \cdots & v^1 w^n \\ \vdots & & \vdots \\ v^m w^1{}_1 & \cdots & v^m w^n \end{pmatrix} = \mathrm{v}^i \, \mathrm{w}^j \, e_i \otimes f_j$$

Note. The basis vectors e_i and f_j may seem to violate the superscript convention for column vectors. The subscripts are necessary to express $\mathrm{v}^i \, \mathrm{w}^j \, e_i \otimes f_j$ in Einstein notation. To make the convention work, think of

$$e_i = \begin{pmatrix} e_i^{\,1} \\ \vdots \\ e_i^{\,i} \\ \vdots \\ e_i^{\,m} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \text{ and } f_j = \begin{pmatrix} f_j^{\,1} \\ \vdots \\ f_j^{\,j} \\ \vdots \\ f_j^{\,m} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}.$$

Definition. The **tensor product of vector spaces V and W** is the vector space $\mathbf{V} \otimes \mathbf{W} : \left\{ \alpha^{ij} \, e_i \otimes f_j : \alpha^{ij} \text{ is a scalar} \right\}$ with addition defined component-wise:

$$P^{ij} + Q^{ij} = \begin{pmatrix} \alpha^{11} & \cdots & \alpha^{1n} \\ \vdots & & \vdots \\ \alpha^{m1} & \cdots & \alpha^{mn} \end{pmatrix} + \begin{pmatrix} \beta^{11} & \cdots & \beta^{1n} \\ \vdots & & \vdots \\ \beta^{m1} & \cdots & \beta^{mn} \end{pmatrix} = \begin{pmatrix} \alpha^{11} + \beta^{11} & \cdots & \alpha^{1n} + \beta^{1n} \\ \vdots & & \vdots \\ \alpha^{m1} + \beta^{m1} & \cdots & \alpha^{mn} + \beta^{mn} \end{pmatrix}.$$

The additive identity is $0 = 0 \otimes w = v \otimes 0 = 0 \otimes 0$ and the additive inverse is $-\left(v \otimes w\right) = \left(-v\right) \otimes w = v \otimes \left(-w\right)$.

The definitions of zero and additive inverse are consistent with the Bilinear Product Rules on page 28:

- $\mathbf{0} = \mathbf{0} \otimes \boldsymbol{w} = \boldsymbol{v} \otimes \mathbf{0}$ follows from (1) by setting $\alpha = 0$. ✔
- Of course, this includes $\mathbf{0} = \mathbf{0} \otimes \mathbf{0}$ since (1) holds for $v = 0 = w$. ✔
- Also, $\mathbf{0} + \boldsymbol{v} \otimes \boldsymbol{w} = 0 \otimes w + v \otimes w \overset{(2)}{=} \left(0 + v\right) \otimes w = \boldsymbol{v} \otimes \boldsymbol{w}$ ✔
- $\boldsymbol{v} \otimes \boldsymbol{w} + \left(-\boldsymbol{v}\right) \otimes \boldsymbol{w} \overset{(2)}{=} \left(v - v\right) \otimes w = \mathbf{0}$ ✔

Note. Dim (V$\otimes$W) = (Dim V) (Dim W):

$\{e_i \otimes f_j\}$ has nm elements and is a basis for V$\otimes$W.

Example. $Q=\begin{bmatrix} 0 & 11 \\ -11 & 0 \end{bmatrix}$ is an example of an element of V$\otimes$W that cannot be expressed as $v\otimes w$ for any v$\in$V and $w\in$W:

Suppose $v=\begin{bmatrix} a \\ b \end{bmatrix}$ and $w=\begin{bmatrix} r \\ s \end{bmatrix}$. Then $v\otimes w=\begin{bmatrix} ar & as \\ br & bs \end{bmatrix}$.

So a = 0 or r = 0. If a = 0 then

$$v\otimes w=\begin{bmatrix} 0 & 0 \\ br & bs \end{bmatrix}\neq Q.$$

If r = 0 then

$$v\otimes w=\begin{bmatrix} 0 & as \\ 0 & bs \end{bmatrix}\neq Q.$$

Theorem. Suppose $x\otimes y=v\otimes w$ where $x=x^i e_i$, $y=y^j e_j$, $v=v^i e_i$, and $w=w^j e_j$. Then $\forall i,j\;\; x^i y^j=v^i w^j$.

Proof. Since $\otimes$ is a bilinear operation,

$$0=x\otimes y-v\otimes w=x^i e_i\otimes y^j f_j - v^i e_i\otimes w^j f_j$$
$$=x^i y^j e_i\otimes f_j - v^i w^j e_i\otimes f_j=\left(x^i y^j - v^i w^j\right)e_i\otimes f_j \quad \blacksquare$$

Theorem. If $x=\alpha v$ and $y=\frac{1}{\alpha}w$, then $x\otimes y=v\otimes w$.

Proof: $x^i y^j=\alpha v^i \frac{1}{\alpha} w^j = v^i w^j$ for all i, j. $\blacksquare$

Notation. $Q^{ab}=x^a\otimes w^b$ denotes $x^a w^b$, the a-b component of the tensor $Q=x\otimes w$. Thus, $Q^{ba}=x^b\otimes w^a$ denotes $x^b w^a$, the b-a component. When there is no confusion we often refer to the tensor Q as Q^{ab}.

Definition. Let S be a linear transformation on V and T a linear transformation on W. The **tensor product of S and T** is the bilinear transformation

$$S\otimes T: \mathrm{V}\otimes\mathrm{W}\to\mathrm{V}\otimes\mathrm{W}: S\otimes T\left(v\otimes w\right)=Sv\otimes Tw.$$

Note. In matrix notation, $S = S^a{}_b$ and $T = T^c{}_d$ are 2-dimensional arrays, i.e., n x n matrices. Since we don't mix S and T, $S \otimes T = R^{ac}_{bd}$, a 4-dimensional array, or an n x n x n x n matrix. Similarly, $v \otimes w = u^{ij}$ is a 2-dimensional array.

Definition. Let V be an n-dimensional vector space and V* its dual space. A **tensor product space of V** is $\mathcal{V} = \mathbf{V}_a^* \otimes \cdots \otimes \mathbf{V}_c^* \otimes \mathbf{V}^f \otimes \cdots \otimes \mathbf{V}^h$ where p and q are positive integers, V_a, …, V_c are q copies of V*, and V^f, …, V^h are p copies of V.

Definition. **An element of $\mathcal{V}$** can be denoted

$$\boldsymbol{Q} = \boldsymbol{Q}^{f \cdots h}_{a \cdots c} = y_a \otimes \cdots \otimes y_c \otimes x^f \otimes \cdots \otimes x^h .$$

Recall from Chapter 12 that ***Q* is a $\begin{bmatrix} p \\ q \end{bmatrix}$-valent tensor over V**, an **abstract quantity** with p upper and q lower indices. Recall that "abstract" means that Q is not tied to a particular basis for V.

Theorem. [13.38] $\mathcal{V}$ is an n^{p+q}-dimensional vector space.
Proof. Recall that if a vector space V is composed of column vectors then the dual space V^* can be considered to be composed of row vectors.

Let { e } be a basis for V. Then $\{f\}$ is a basis for $V^{\mathsf{T}} \cong V^*$ where $f = e^{\mathsf{T}}$. Let $\{f_a\}, \cdots, \{f_c\}$ be q copies of { f } and $\{e^f\}, \cdots, \{e^h\}$ be p copies of { e }. Then $\{f_a \otimes \cdots \otimes f_c \otimes e^f \otimes \cdots \otimes e^h\}$ is a basis for $\mathcal{V}$. Since the basis contains (q+p) factors with n choices for each factor, Dim $\mathcal{V} = n^{p+q}$.

The vector space properties hold for $\mathcal{V}$ because $\otimes \cdots \otimes$ is multilinear. For example, $0 = 0^{f \cdots h}_{a \cdots c} = 0_a \otimes \cdots \otimes 0_c \otimes 0^f \otimes \cdots \otimes 0^h$ is the (additive) group identity of $\mathcal{V}$:

$$\begin{aligned} 0 + Q^{f \cdots h}_{a \cdots c} &= \left(0_a \otimes \cdots \otimes 0_c \otimes 0^f \otimes \cdots \otimes 0^h\right) + \left(y_a \otimes \cdots \otimes y_c \otimes x^f \otimes \cdots \otimes x^h\right) \\ &= \left(0_a + y_a\right) \otimes \cdots \otimes \left(0_c + y_c\right) \otimes \left(0^f + x^f\right) \otimes \cdots \otimes \left(0^h + x^h\right) \\ &= y_a \otimes \cdots \otimes y_c \otimes x^f \otimes \cdots \otimes x^h = Q^{f \cdots h}_{a \cdots c} \end{aligned}$$

and similarly $Q^{f \cdots h}_{a \cdots c} + 0 = Q^{f \cdots h}_{a \cdots c}$.

The additive inverse of $Q^{f \cdots h}_{a \cdots c}$ is $-Q^{f \cdots h}_{a \cdots c}$.

If $P^{f \cdots h}_{a \cdots c} = z_a \otimes \cdots \otimes z_c \otimes w^f \otimes \cdots \otimes w^h$ then

$$\begin{aligned}P+Q&=\left(z_a\otimes\cdots\otimes z_c\otimes w^f\otimes\cdots\otimes w^h\right)+\left(y_a\otimes\cdots\otimes y_c\otimes x^f\otimes\cdots\otimes x^h\right)\\&=\left(z_a+y_a\right)\otimes\cdots\otimes\left(z_a+y_a\right)\otimes\left(w^f+x^f\right)\otimes\cdots\otimes\left(w^h+x^h\right)\\&=\left(y_a+z_a\right)\otimes\cdots\otimes\left(y_a+z_a\right)\otimes\left(x^f+w^f\right)\otimes\cdots\otimes\left(x^h+w^h\right)\\&=\left(y_a\otimes\cdots\otimes y_c\otimes x^f\otimes\cdots\otimes x^h\right)+\left(z_a\otimes\cdots\otimes z_c\otimes w^f\otimes\cdots\otimes w^h\right)\\&=Q+P.\end{aligned}$$

As one last example, if α is a scalar then

$$\begin{aligned}\alpha\left(P+Q\right)&=\alpha\left(z_a+y_a\right)\otimes\cdots\otimes\alpha\left(z_a+y_a\right)\otimes\alpha\left(w^f+x^f\right)\otimes\cdots\otimes\alpha\left(w^h+x^h\right)\\&=\left(\alpha z_a\otimes\cdots\otimes\alpha z_c\otimes\alpha w^f\otimes\cdots\otimes\alpha w^h\right)+\left(\alpha y_a\otimes\cdots\otimes\alpha y_c\otimes\alpha x^f\otimes\cdots\otimes\alpha x^h\right)\\&=\alpha P+\alpha Q\end{aligned}$$

■

$Q\in\mathcal{V}$ can be expressed as an $n\times n\times\cdots\times n$ generalized matrix, a **(*p*+*q*)-dimensional array**. For example, Q_a^f**,** Q_{ab}**,** and Q^{fg} are *n* x *n* matrices, 2-dimensional arrays. Q_c^{ab} is an *n* x *n* x *n* matrix, a 3-dimensional array, a cube.

Array Dimension	Tensor	# of Entries
0	Scalar	$n^0 = 1$
1	Vector	n
2	Matrix	n^2
3	3-Tensor (cube of numbers)	n^3
k	*k*-Tensor (*k*-dimensional hypercube of numbers)	n^k

Q$^{\mathrm{T}}$ **can be considered to be an element of** $\mathcal{V}$***.** Let

$$Q_{a\cdots c}^{f\cdots h}=y_a\otimes\cdots\otimes y_c\otimes x^f\otimes\cdots\otimes x^h$$

and

$$P_{a\cdots c}^{f\cdots h}=z_a\otimes\cdots\otimes z_c\otimes w^f\otimes\cdots\otimes w^h\in\mathcal{V}.$$

Then

$$Q^{\mathrm{T}}:\mathcal{V}\to\mathbb{R}\text{ or }\mathbb{C}:Q^{\mathrm{T}}\left(P\right)=z_a\cdots z_c\,Q_{f\cdots h}^{a\cdots c}\,w^f\cdots w^h.$$

Observe that $w^f\cdots w^h$ is a column tensor, so $Q_{f\cdots h}^{a\cdots c}\,w^f\cdots w^h$ is also a column tensor. $z_a\cdots z_c$ is a row tensor, so $z_a\cdots z_c\left(Q_{f\cdots h}^{a\cdots c}\,w^f\cdots w^h\right)$ is a row tensor times a column tensor, yielding a scalar. Were we simply to multiply $Q^{a\cdots c}_{f\cdots h}\,P^{f\cdots h}_{a\cdots c}$ we would end up with a *q* x *q* matrix, not a scalar as desired.

Example. Let V = $\mathbb{R}^2$, Let *S* and *T* be linear transformations on V. Given a basis for V, *S* and *T* can be represented by matrices *A* and *B*, respectively:

$$A^a{}_{a'} = \begin{pmatrix} a^1{}_1 & a^1{}_2 \\ a^2{}_1 & b^2{}_2 \end{pmatrix} \text{ and } B^b{}_{b'} = \begin{pmatrix} b^1{}_1 & b^1{}_2 \\ b^2{}_1 & b^2{}_2 \end{pmatrix}.$$

Then

$$A^a{}_{a'} \otimes B^b{}_{b'} = \begin{bmatrix} a^1{}_1 \begin{pmatrix} b^1{}_1 & b^1{}_2 \\ b^2{}_1 & b^2{}_2 \end{pmatrix} & a^1{}_2 \begin{pmatrix} b^1{}_1 & b^1{}_2 \\ b^2{}_1 & b^2{}_2 \end{pmatrix} \\ a^2{}_1 \begin{pmatrix} b^1{}_1 & b^1{}_2 \\ b^2{}_1 & b^2{}_2 \end{pmatrix} & a^2{}_2 \begin{pmatrix} b^1{}_1 & b^1{}_2 \\ b^2{}_1 & b^2{}_2 \end{pmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{pmatrix} a^1{}_1 b^1{}_1 & a^1{}_1 b^1{}_2 \\ a^1{}_1 b^2{}_1 & a^1{}_1 b^2{}_2 \end{pmatrix} & \begin{pmatrix} a^1{}_2 b^1{}_1 & a^1{}_2 b^1{}_2 \\ a^1{}_2 b^2{}_1 & a^1{}_2 b^2{}_2 \end{pmatrix} \\ \begin{pmatrix} a^2{}_1 b^1{}_1 & a^2{}_1 b^1{}_2 \\ a^2{}_1 b^2{}_1 & a^2{}_1 b^2{}_2 \end{pmatrix} & \begin{pmatrix} a^2{}_2 b^1{}_1 & a^2{}_2 b^1{}_2 \\ a^2{}_2 b^2{}_1 & a^2{}_2 b^2{}_2 \end{pmatrix} \end{bmatrix},$$

a 4-dimensional hypercube. That is, a 2x2 tensor (like *A* or *B*) is a matrix, a 2-D square with 4 entries. A 2x2x2 tensor is a 3-D cube with 8 entries. *C* = $A \otimes B$ is a 2x2x2x2 tensor, a 4-D hypercube with 16 entries.

If $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ and $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ are vectors, then

$$\begin{aligned} A^a{}_{a'} \otimes B^b{}_{b'} \left(v^{a'} \otimes w^{b'} \right) &= A^a{}_{a'} \left(v^{a'} \right) \otimes B^b{}_{b'} \left(w^{b'} \right) \\ &= \begin{pmatrix} a^1{}_1 & a^1{}_2 \\ a^2{}_1 & a^2{}_2 \end{pmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \otimes \begin{pmatrix} b^1{}_1 & b^1{}_2 \\ b^2{}_1 & b^2{}_2 \end{pmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \\ &= \begin{pmatrix} a^1{}_1 v_1 + a^1{}_2 v_2 \\ a^2{}_1 v_1 + a^2{}_2 v_2 \end{pmatrix} \otimes \begin{pmatrix} b^1{}_1 w_1 + b^1{}_2 w_2 \\ b^2{}_1 w_1 + b^2{}_2 w_2 \end{pmatrix}. \end{aligned}$$

Remember, we don't multiply this out because we don't mix *v* and *w.* However, we can further expand this to

$$= \begin{pmatrix} \left(a^1{}_1v_1 + a^1{}_2v_2\right) \otimes \begin{pmatrix} b^1{}_1w_1 + b^1{}_2w_2 \\ b^2{}_1w_1 + b^2{}_2w_2 \end{pmatrix} \\ \\ \left(a^2{}_1v_1 + a^2{}_2v_2\right) \otimes \begin{pmatrix} b^1{}_1w_1 + b^1{}_2w_2 \\ b^2{}_1w_1 + b^2{}_2w_2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \left(a^1{}_1v_1 + a^1{}_2v_2\right)\left(b^1{}_1w_1 + b^1{}_2w_2\right) \\ \left(a^1{}_1v_1 + a^1{}_2v_2\right)\left(b^2{}_1w_1 + b^2{}_2w_2\right) \\ \left(a^2{}_1v_1 + a^2{}_2v_2\right)\left(b^1{}_1w_1 + b^1{}_2w_2\right) \\ \left(a^2{}_1v_1 + a^2{}_2v_2\right)\left(b^2{}_1w_1 + b^2{}_2w_2\right) \end{pmatrix}$$

Theorem. [13.39] The linear transformation $x \mapsto Tx$ (or $x^a \mapsto T^a{}_b x^b$) on V induces a linear transformation $\mathcal{T}$: $Q^{f\cdots h}_{a\cdots c} \mapsto S^{a'}{}_a \cdots S^{c'}{}_c\, T^f{}_{f'} \cdots T^h{}_{h'}\, Q^{f'\cdots h'}_{a'\cdots c'}$ on $\mathcal{V}$ where $S = \left(T^{-1}\right)^{\mathsf{T}}$ is the transpose of T^{-1}.

Proof.
We must first show that $\mathcal{T}\left(Q\right) \in \mathcal{V}$:

Let $Q^{f\cdots h}_{a\cdots c} = y_a \otimes \cdots \otimes y_c \otimes x^f \otimes \cdots \otimes x^h$ where $y_a, \cdots, y_c \in V^{\mathsf{T}}$ and $x^f, \cdots, x^h \in V$.
Recall that $S^{a'}{}_a \cdots S^{c'}{}_c T^f{}_{f'} \cdots T^h{}_{h'} = S^{a'}{}_a \otimes \cdots \otimes S^{c'}{}_c \otimes T^f{}_{f'} \otimes \cdots \otimes T^h{}_{h'}$. Therefore

$$\begin{aligned} \mathcal{T}\left(Q\right) &= S^{a'}{}_a \otimes \cdots \otimes S^{c'}{}_c \otimes T^f{}_{f'} \otimes \cdots \otimes T^h{}_{h'} \left(y_{a'} \otimes \cdots \otimes y_{c'} \otimes x^{f'} \otimes \cdots \otimes x^{h'}\right) \\ &= S^{a'}{}_a y_{a'} \otimes \cdots \otimes S^{c'}{}_c y_{c'} \otimes T^f{}_{f'} x^{f'} \otimes \cdots \otimes T^h{}_{h'} x^{h'}. \end{aligned}$$

Also, $S : V^{\mathsf{T}} \to V^{\mathsf{T}}$ and $T : V \to V$. So, for example, $S^{a'}{}_a y_{a'}$ is a sum of vectors in V^{T}. Therefore

$$S^{a'}{}_a y_{a'} \in V^{\mathsf{T}} \cong V^*,$$

and thus

$$\mathcal{T}\left(Q\right) \in V^* \otimes \cdots \otimes V^* \otimes V \otimes \cdots \otimes V = \mathcal{V} \qquad ✔$$

To show that $\mathcal{T}$ is linear, let *P* and *Q* be $\begin{bmatrix} p \\ q \end{bmatrix}$-valent tensors, α a scalar, and *R* = *P* + *Q*. Because the tensor product is multilinear,

$$\begin{aligned} \mathcal{T}\left(P^{f\cdots h}_{a\cdots c} + Q^{f\cdots h}_{a\cdots c}\right) = \mathcal{T}\left(R^{f\cdots h}_{a\cdots c}\right) &= S^{a'}{}_a \cdots S^{c'}{}_c\, T^f{}_{f'} \cdots T^h{}_{h'}\, R^{f'\cdots h'}_{a'\cdots c'} \\ &= S^{a'}{}_a \cdots S^{c'}{}_c\, T^f{}_{f'} \cdots T^h{}_{h'} \left(P^{f'\cdots h'}_{a'\cdots c'} + Q^{f'\cdots h'}_{a'\cdots c'}\right) \\ &= S^{a'}{}_a \cdots S^{c'}{}_c\, T^f{}_{f'} \cdots T^h{}_{h'}\, P^{f'\cdots h'}_{a'\cdots c'} + S^{a'}{}_a \cdots S^{c'}{}_c\, T^f{}_{f'} \cdots T^h{}_{h'}\, Q^{f'\cdots h'}_{a'\cdots c'} \\ &= \mathcal{T}\left(P^{f\cdots h}_{a\cdots c}\right) + \mathcal{T}\left(Q^{f\cdots h}_{a\cdots c}\right) \end{aligned}$$

and

$$T\left(\alpha\, Q^{f\cdots h}_{a\cdots c}\right) = T\left(\left(\alpha\, Q\right)^{f\cdots h}_{a\cdots c}\right) = S^{a}{}_{a'}\cdots S^{c}{}_{c'}\, T^{f}{}_{f'}\cdots T^{h}{}_{h'}\left(\alpha\, Q\right)^{f'\cdots h'}_{a'\cdots c'}$$

$$= \alpha\, S^{a}{}_{a'}\cdots S^{c'}{}_{c'}\, T^{f}{}_{f'}\cdots T^{h}{}_{h'} Q^{f'\cdots h'}_{a'\cdots c'}$$

$$= \alpha\, T\left(Q^{f\cdots h}_{a\cdots c}\right) \qquad \blacksquare$$

The next lemma shows that the multilinear tensor product definition enables a certain amount of tensor interchanging even though there is no commutativity per se.

Lemma. Let V be a vector space and $T: \mathrm{V} \to \mathrm{V}$ a linear transformation. Let $S = T^{-1}$. Then

$$S^{a'}{}_{a}\cdots S^{c'}{}_{c}\, T^{f}{}_{f'}\cdots T^{h}{}_{h'}\, S^{a''}{}_{a'}\cdots S^{c''}{}_{c'}\, T^{f'}{}_{f''}\cdots T^{h'}{}_{h''}$$

$$= S^{a'}{}_{a}\, S^{a''}{}_{a'}\cdots S^{c'}{}_{c}\, S^{c''}{}_{c'}\, T^{f}{}_{f'}\, T^{f'}{}_{f''}\cdots T^{h}{}_{h'}\, T^{h'}{}_{h''}\,.$$

$$\begin{tikzcd} \mathcal{V} \arrow[r, "T_2"] \arrow[dr, dashed, "T_{12}"'] & \mathcal{V} \arrow[d, "T_1"] \\ & \mathcal{V} \end{tikzcd}$$

Proof. The lemma is summarized in the figure at the right. $\mathcal{V} = \mathrm{V}^* \otimes \cdots \otimes \mathrm{V}^* \otimes \mathrm{V} \otimes \cdots \otimes \mathrm{V}$, $T_1\, T_2$ is represented by the first expression, and T_{12} is represented by the second one. Let $Q^{f\cdots h}_{a\cdots c} = y_a \otimes \cdots \otimes y_c \otimes x^f \otimes \cdots \otimes x^h \in \mathcal{V}$. We wish to show that $T_1\, T_2 Q = T_{12} Q$, which we get by applying the definition of the (multilinear) tensor product twice:

$$S^{a'}{}_{a}\cdots S^{c'}{}_{c}\, T^{f}{}_{f'}\cdots T^{h}{}_{h'}\, S^{a''}{}_{a'}\cdots S^{c''}{}_{c'}\, T^{f'}{}_{f''}\cdots T^{h'}{}_{h''}\, Q^{f''\cdots h''}_{a''\cdots c''}$$

$$= S^{a'}{}_{a} \otimes\cdots\otimes S^{c'}{}_{c} \otimes T^{f}{}_{f'} \otimes\cdots\otimes T^{h}{}_{h'} \otimes S^{a''}{}_{a'} \otimes\cdots\otimes S^{c''}{}_{c'} \otimes T^{f'}{}_{f''} \otimes\cdots\otimes T^{h'}{}_{h''}$$
$$\left(y_{a''} \otimes\cdots\otimes y_{c''} \otimes x^{f''} \otimes\cdots\otimes x^{h''}\right)$$

$$= S^{a'}{}_{a} \otimes\cdots\otimes S^{c'}{}_{c} \otimes T^{f}{}_{f'} \otimes\cdots\otimes T^{h}{}_{h'}\left(S^{a''}{}_{a'} y_{a''} \otimes\cdots\otimes S^{c''}{}_{c'} y_{c''} \otimes T^{f'}{}_{f''} x^{f''} \otimes\cdots\otimes T^{h'}{}_{h''} x^{h''}\right)$$

$$= S^{a'}{}_{a} S^{a''}{}_{a'} y_{a''} \otimes\cdots\otimes S^{c'}{}_{c} S^{c''}{}_{c'} y_{c''} \otimes T^{f}{}_{f'} T^{f'}{}_{f''} x^{f''} \otimes\cdots\otimes T^{h}{}_{h'} T^{h'}{}_{h''} x^{h''}$$

$$= S^{a'}{}_{a}\, S^{a''}{}_{a'}\cdots S^{c'}{}_{c}\, S^{c''}{}_{c'}\, T^{f}{}_{f'}\, T^{f'}{}_{f''}\cdots T^{h}{}_{h'}\, T^{h'}{}_{h''}\, Q^{f''\cdots h''}_{a''\cdots c''} \qquad \blacksquare$$

In the next theorem, the mapping $y \mapsto y\, S^{\mathsf{T}}$ plays the central role, replacing $y \mapsto y\, S$ that was used to show that V* is a representation space. Since S is a square matrix, S^{T} (as well as S) is a linear transformation on V and $y\, S^{\mathsf{T}} \in \mathrm{V}^*$ as the figure at the right shows.

$$\begin{tikzcd} \mathrm{V} \arrow[r, "S^{\mathsf{T}}"] \arrow[dr, dashed, "y S^{\mathsf{T}}"'] & \mathrm{V} \arrow[d, "y"] \\ & \mathbb{R} \text{ or } \mathbb{C} \end{tikzcd}$$

Theorem. If V is a representation space for a group G, then so is the tensor product space $\mathcal{V}$.

Proof. By definition of representation, there is a subgroup $\mathcal{G} \subset \mathrm{GL}(n)$ and an isomorphism $T : G \to \mathcal{G}$ such that $\forall g \in G \;\; T(g) : \mathrm{V} \to \mathrm{V}$ is a linear transformation on V. We seek another subgroup $\mathcal{G}^* \subset \mathrm{GL}(n)$ and an isomorphism $\mathcal{T} : G \to \mathcal{G}^*$ such that $\forall g \in G \;\; \mathcal{T}(g) : \mathcal{V} \to \mathcal{V}$ is a linear transformation on $\mathcal{V}$. Denote

(1) $T_g = T(g)$ for $g \in$ G.

T_g is a linear transformation on V. Since T is an isomorphism defined on G,

(2) $T_{g_1 g_2} \overset{(1)}{=} T(g_1 g_2) = T(g_1)\,T(g_2) \overset{(1)}{=} T_{g_1} T_{g_2}$ for $g_1, g_2 \in G$.

Set

(3) $S_g = T_g^{-1}$.

By prior theorem [13.39], T_g induces a linear transformation $\mathcal{T}_g$ on $\mathcal{V}$:

For $Q^{f \cdots h}_{a \cdots c} \in \mathcal{V}$, define

$$\mathcal{T}_g : \mathcal{V} \to \mathcal{V} : \mathcal{T}_g\left(Q^{f \cdots h}_{a \cdots c}\right) = \left(S^{\mathrm{T}}\right)^{a'}{}_{a} \otimes \cdots \otimes \left(S^{\mathrm{T}}\right)^{c'}{}_{c} \otimes T^{f}{}_{f'} \otimes \cdots \otimes T^{h}{}_{h'} \left(Q^{f' \cdots h'}_{a' \cdots c'}\right)$$

or, in abbreviated format,

(4) $\mathcal{T}_g = \left(S^{\mathrm{T}}\right)^{a'}{}_{a} \cdots \left(S^{\mathrm{T}}\right)^{c'}{}_{c}\, T^{f}{}_{f'} \cdots T^{h}{}_{h'}$.

Define

$\mathcal{G}^* = \left\{\mathcal{T}_g : g \in G\right\}$

and

(5) $\mathcal{T} : G \to \mathcal{G}^* : \mathcal{T}(g) = \mathcal{T}_g$ for $g \in$ G.

$\mathcal{G}^*$ is clearly a subgroup of GL(n). We show that $\mathcal{T}$ is a homomorphism. Let $g_1, g_2 \in$ G and set

(6) $g_3 = g_1 g_2$.

We need to show that $\mathcal{T}(g_1 g_2) = \mathcal{T}(g_1)\,\mathcal{T}(g_2)$.

(7) $T_{g_3} \overset{(6)}{=} T_{g_1 g_2} \overset{(2)}{=} T_{g_1} T_{g_2}$ or $T_{g_3}{}^{f}{}_{f'} = T_{g_1}{}^{f}{}_{f'} T_{g_2}{}^{f'}{}_{f''}$

$\Rightarrow \qquad S_{g_3} \overset{(3)}{=} T_{g_3}{}^{-1} \overset{(7)}{=} T_{g_2}{}^{-1} T_{g_1}{}^{-1} \overset{(3)}{=} S_{g_2} S_{g_1}$.

Therefore

(8) $S_{g_3}^{\mathsf{T}} = S_{g_1}^{\mathsf{T}} S_{g_2}^{\mathsf{T}}$ or $\left(S_{g_3}{}^{\mathsf{T}}\right)^{a'}{}_{a} = \left(S_{g_1}{}^{\mathsf{T}}\right)^{a'}{}_{a} \left(S_{g_2}{}^{\mathsf{T}}\right)^{a''}{}_{a'}$.

Observe that the inverse operation changed the order of g_1 and g_2. Then the transpose operation changed it back to the desired order. So

$$(9) \quad \mathcal{T}\left(g_1 g_2\right) \overset{(6)}{=} \mathcal{T}\left(g_3\right) \overset{(5)}{=} \mathcal{T}_{g_3} \overset{(4)}{=} \left(S_{g_3}{}^{\mathsf{T}}\right)^{a'}{}_{a} \cdots \left(S_{g_3}{}^{\mathsf{T}}\right)^{c'}{}_{c} T_{g_3}{}^{f}{}_{f'} \cdots T_{g_3}{}^{h}{}_{h'}$$
$$\overset{(7,\,8)}{=} \left(S_{g_1}{}^{\mathsf{T}}\right)^{a'}{}_{a} \left(S_{g_2}{}^{\mathsf{T}}\right)^{a''}{}_{a'} \cdots \left(S_{g_1}{}^{\mathsf{T}}\right)^{c'}{}_{c} \left(S_{g_2}{}^{\mathsf{T}}\right)^{c''}{}_{c'} T_{g_1}{}^{f}{}_{f'} T_{g_2}{}^{f'}{}_{f''} \cdots T_{g_1}{}^{h}{}_{h'} T_{g_2}{}^{h'}{}_{h''}$$

and

$$(10) \quad \mathcal{T}\left(g_1\right)\mathcal{T}\left(g_2\right) \overset{(5)}{=} \mathcal{T}_{g_1}\mathcal{T}_{g_2}$$
$$\overset{(4)}{=} \left(S_{g_1}{}^{\mathsf{T}}\right)^{a}{}_{a'} \cdots \left(S_{g_1}{}^{\mathsf{T}}\right)^{c}{}_{c'} T_{g_1}{}^{f}{}_{f'} \cdots T_{g_1}{}^{h}{}_{h'} \left(S_{g_2}{}^{\mathsf{T}}\right)^{a''}{}_{a'} \cdots \left(S_{g_2}{}^{\mathsf{T}}\right)^{c''}{}_{c'} T_{g_2}{}^{f'}{}_{f''} \cdots T_{g_2}{}^{h'}{}_{h''}$$

By the lemma, (9) = (10) and hence $\mathcal{T}$ is a homomorphism. ✔

To show that $\mathcal{T}$ is an isomorphism, we must show that it is 1-1. Again, let $g_1, g_2 \in \mathrm{G}$. We must show that $\mathcal{T}\left(g_1\right) = \mathcal{T}\left(g_2\right) \quad \Rightarrow \quad g_1 = g_2$. To simplify notation, set

$$T = \mathcal{T}\left(g_1\right)$$

and

$$N = \mathcal{T}\left(g_2\right).$$

T and *N* are linear transformations on V. Set

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

and

$M = N^{-1}$.

By (4),

$$T(g_1) \overset{(5)}{=} T_{g_1} \overset{(4)}{=} \left(S^{\mathrm{T}}\right)^{a'}{}_{a} \cdots \left(S^{\mathrm{T}}\right)^{c'}{}_{c}\, T^{f}{}_{f'} \cdots T^{h}{}_{h'}$$

and

$$T(g_2) \overset{(5)}{=} T_{g_2} \overset{(4)}{=} \left(M^{\mathrm{T}}\right)^{a'}{}_{a} \cdots \left(M^{\mathrm{T}}\right)^{c'}{}_{c}\, N^{f}{}_{f'} \cdots N^{h}{}_{h'}.$$

So,

$$T(g_1) = T(g_2) \quad \Leftrightarrow \quad S^{a'}{}_{a} = M^{a'}{}_{a}, \cdots, S^{c'}{}_{c} = M^{c'}{}_{c}, T^{f}{}_{f'} = N^{f}{}_{f'}, \cdots, T^{h}{}_{h'} = N^{h}{}_{h'}$$

because we don't mix dissimilar indices. Each of these expressions is equivalent to $T = N$. So,

$$T(g_1) = T(g_2) \quad \Leftrightarrow \quad T = N \quad \Leftrightarrow \quad T(g_1) = T(g_2) \quad \Rightarrow \quad g_1 = g_2$$

because T is an isomorphism. ■

Definition. Let V be an n-dimensional vector space and $Q^{f\cdots h} \in \mathcal{V} = \mathrm{V} \otimes \cdots \otimes \mathrm{V}$ a $\begin{bmatrix} p \\ 0 \end{bmatrix}$-valent tensor. The **symmetric part of Q** is $Q^{(f\cdots h)} = \frac{1}{p!}\sum_{\pi} Q^{\pi(f)\cdots\pi(h)} = \frac{1}{2}\left(Q^{fg\cdots h} + Q^{gf\cdots h}\right)$ and the **antisymmetric part of Q** is $Q^{[f\cdots h]} = \frac{1}{p!}\sum_{\pi} \mathrm{Sign}(\pi)\, Q^{\pi(f)\cdots\pi(h)} = \frac{1}{2}\left(Q^{fg\cdots h} - Q^{gf\cdots h}\right)$. Notice that $Q = Q^{(f\cdots h)} + Q^{[f\cdots h]}$. The **symmetric space** is $\mathcal{V}_+ = \left\{Q^{(f\cdots h)} : Q \in \mathcal{V}\right\}$ and the **antisymmetric space** is $\mathcal{V}_- = \left\{Q^{[f\cdots h]} : Q \in \mathcal{V}\right\}$.

Theorem. [13.40]

(1) $\mathcal{V}_+$ and $\mathcal{V}_-$ are vector spaces

(2) $\mathcal{V}_+ \cap \mathcal{V}_- = \{0\}$

(3) $\mathcal{V} = \mathcal{V}_+ \oplus \mathcal{V}_-$ (i.e., $\mathcal{V} = \mathcal{V}_+ + \mathcal{V}_-$ and $\mathcal{V}_+ \cap \mathcal{V}_- = \{0\}$)

(4) Dim $\mathcal{V}_+ = \frac{n}{2}(n+1)$ and Dim $\mathcal{V}_- = \frac{n}{2}(n-1)$

(5) $P \in \mathcal{V}_+ \Rightarrow P^{f\cdots h} = P^{\pi(f)\cdots\pi(h)}$ for any permutation π,

(6) $P \in \mathcal{V}_- \Rightarrow P^{f \cdots h} = \begin{cases} P^{\pi(f) \cdots \pi(h)} & \text{if } \pi \text{ is even} \\ -P^{\pi(f) \cdots \pi(h)} & \text{if } \pi \text{ is odd} \end{cases}$

Note. There is a parallel theory for $\begin{bmatrix} 0 \\ q \end{bmatrix}$-valent tensors $Q_{a \cdots c} \in \mathrm{V}^* \otimes \cdots \otimes \mathrm{V}^*$.

Example. Let $Q \in \mathcal{V} = \mathrm{V} \otimes \mathrm{V}$ and $P^{ab} = Q^{(ab)} \in \mathcal{V}_+$. Then $P^{ba} = P^{ab}$. Let $R^{ab} = Q^{[ab]} \in \mathcal{V}_-$. Then $R^{ba} = -R^{ab}$.