3. Lie Derivatives and Lie Groups

3.1. Introduction: How a Vector Field Maps a Manifold into Itself

As mentioned in §2.12, the (non-intersecting) integral curves of a vector field $\bar{V} = \frac{d}{d\lambda}$ can fill up a region *U* of a manifold *M* and form a **congruence**. In particular, each point in *U* is passed by one and only one curve. Since each curve is a 1-D manifold, the congruence is (*n*–1)-D manifold. Now, by increasing the parameter λ by $\Delta\lambda$, every point in *M* will be mapped to a point a "distance" $\Delta\lambda$ down the curve it is on. This is a 1-1 map of *M* into itself if $\bar{V}$ is sufficiently well behaved (a C^1 field will do). If $\bar{V}$ is C^∞, the map is a **diffeomorphism**. If the map exists for all $\Delta\lambda$, there will be a 1-D differentiable family of them, which form a 1-parameter Lie group. Such a mapping is calle a **dragging** along the congruence, or a **Lie dragging**.

3.2. Lie Dragging a Function

Lie dragging a function *f* on *M* along the congruence of parameter λ by $\Delta\lambda$ produces a new function $f^*_{\Delta\lambda}$ defined by

$$f^*_{\Delta\lambda}(Q) = f(P)$$

where *Q* is the point reached by dragging *P*. If it happens that

$$f^*_{\Delta\lambda}(Q) = f(Q) \qquad \text{for all } Q$$

then *f* is invariant undering the Lie dragging, i.e.,

$$f^*_{\Delta\lambda} = f$$

If *f* is invariant under draggings of all $\Delta\lambda$, it is called **Lie dragged**. Obviously, the condition for this is that *f* is a constant along each curve in the congruence, i.e., $\frac{df}{d\lambda} = 0$.

3.3. Lie Dragging a Vector Field

The Lie dragging of a vector field $\frac{d}{d\mu}$ along the congruence of $\frac{d}{d\lambda}$ by an amount $\Delta\lambda$ produces a new vector field $\frac{d}{d\mu^*_{\Delta\lambda}}$. Geometrically, this is accomplished by dragging each point on every integral curve $C(\mu)$ along the integral curve $C(\lambda)$ that passes through it by an amount $\Delta\lambda$. The dragged points then form a new curve $C(\mu^*_{\Delta\lambda})$. [see Fig.3.2]

If it happens that $\frac{d}{d\mu^*_{\Delta\lambda}} = \frac{d}{d\mu}$ everywhere, then $\frac{d}{d\mu}$ is said to be **invariant** under the Lie dragging. If it is invariant for all $\Delta\lambda$, then $\frac{d}{d\mu}$ is said to be **Lie dragged** by $\frac{d}{d\lambda}$. Geometrically, this means the integral curves $C(\mu)$ and $C(\lambda)$ form a coordinate grid so that [see §2.14]

$$\left[\frac{d}{d\lambda}, \frac{d}{d\mu}\right] = 0 \qquad (3.1)$$

One can also treat a Lie dragged congruence as generated from a single curve $C_1(\mu)$. Thus, each curve $C(\mu)$ in the congruence is obtained by dragging every point on $C_1(\mu)$ along the congruence of $\frac{d}{d\lambda}$ by a fixed amount. Therefore, λ is constant on each $C(\mu)$. Hence (3.1).

3.4. Lie Derivatives

3.4.0. Preliminary

In defining a derivation of tensors, one must first devise a meaningful way for comparing tensors at different points of the manifold and hence belonging to different tensor spaces. This is accomplished by mapping (moving) one tensor to the tensor space at the other point. Thereafter, standard methods of analysis can be applied since one is dealing only with tensors in the same space.

One way to "move" a vector to another tangent space is to map it to its **parallel** image there. In Riemannian geometry, this involves a definition of **parallelism** and is accomplished by introducing an **affine connection** into the manifold. The resultant derivation is called **covariant derivatives**. [see chapter 6].

Another way for moving tensors that does not involve introducing additional structures to the manifold is by Lie dragging. This gives rise to the **Lie derivatives**.

3.4.1. Lie Derivatives of Scalar Functions

The Lie derivative $\mathcal{L}_{\bar{V}} f$ of a function *f* on *M* at the point $P(\lambda_0)$ on an integral curve of $\bar{V} = \dfrac{d}{d\lambda}$ is defined as

$$\mathcal{L}_{\bar{V}} f\Big|_{\lambda_0} = \lim_{\Delta\lambda\to 0} \frac{f^*_{\Delta\lambda}(\lambda_0) - f(\lambda_0)}{\Delta\lambda}$$

where the function $f^*_{\Delta\lambda}$ is *f* Lie dragged along $\bar{V}$ by $\Delta\lambda$, i.e.,

$$f^*_{\Delta\lambda}(\lambda) = f(\lambda + \Delta\lambda) \qquad \forall\, \lambda$$

on any integral curve of $\bar{V}$. [see §3.2] Thus,

$$\mathcal{L}_{\bar{V}} f\Big|_{\lambda_0} = \lim_{\Delta\lambda\to 0} \frac{f(\lambda_0 + \Delta\lambda) - f(\lambda_0)}{\Delta\lambda} = \frac{df}{d\lambda}\bigg|_{\lambda_0} \tag{3.2}$$

$$= \bar{V}(f)\Big|_{\lambda_0} \tag{3.3}$$

Caution: $\bar{V}(f)$ merely means operating on *f* by the differential operator $\bar{V} = \dfrac{d}{d\lambda}$ so that $\bar{V}(f) = \dfrac{d}{d\lambda}(f) = \dfrac{df}{d\lambda}$. On the other hand,

$$\bar{V}(\tilde{d}f) = \frac{d}{d\lambda}(\tilde{d}f) = \frac{dx^\mu}{d\lambda}\frac{\partial}{\partial x^\mu}\left(\frac{\partial f}{\partial x^\nu}\tilde{d}x^\nu\right) = \frac{dx^\mu}{d\lambda}\frac{\partial f}{\partial x^\nu}\frac{\partial}{\partial x^\mu}\left(\tilde{d}x^\nu\right)$$

$$= \frac{dx^\mu}{d\lambda}\frac{\partial f}{\partial x^\nu}\delta^\nu_\mu = \frac{dx^\mu}{d\lambda}\frac{\partial f}{\partial x^\mu} = \frac{df}{d\lambda}$$

denotes the *contraction* or *inner product* between a vector and a 1-form. In particular,

$$\frac{\partial}{\partial x^\mu}\left(\frac{\partial f}{\partial x^\nu}\tilde{d}x^\nu\right) = \left\langle \bar{e}_\mu, \frac{\partial f}{\partial x^\nu}\tilde{e}^\nu \right\rangle = \frac{\partial f}{\partial x^\nu}\left\langle \bar{e}_\mu, \tilde{e}^\nu \right\rangle = \frac{\partial f}{\partial x^\nu}\delta^\nu_\mu = \frac{\partial f}{\partial x^\nu}\frac{\partial}{\partial x^\mu}\left(\tilde{d}x^\nu\right)$$

3.4.2. Lie Derivatives of Vector Fields

The Lie derivative $\mathcal{L}_{\bar{V}}\bar{U}$ of a vector field $\bar{U}=\dfrac{d}{d\mu}$ along $\bar{V}=\dfrac{d}{d\lambda}$ is defined as

$$\mathcal{L}_{\bar{V}}\bar{U}\Big|_{\lambda_0} = \lim_{\Delta\lambda\to 0}\frac{\bar{U}^*_{\Delta\lambda}\Big|_{\lambda_0}-\bar{U}\Big|_{\lambda_0}}{\Delta\lambda} \qquad (3.5a)$$

so that for an arbitrary function *f* on *M*, we have

$$\left[\mathcal{L}_{\bar{V}}\bar{U}\right](f)\Big|_{\lambda_0} = \lim_{\Delta\lambda\to 0}\left[\frac{\bar{U}^*_{\Delta\lambda}(f)-\bar{U}(f)}{\Delta\lambda}\right]\Bigg|_{\lambda_0} \qquad (3.5)$$

Here, $\bar{U}^*_{\Delta\lambda}=\dfrac{d}{d\mu^*_{\Delta\lambda}}$ is the vector field generated by Lie dragging $\bar{U}\big|_{\lambda_0+\Delta\lambda}$ along $\bar{V}$,

i.e., the vector field $\bar{U}^*_{\Delta\lambda}$ has value

$$\bar{U}^*_{\Delta\lambda}\Big|_{\lambda_0+\Delta\lambda}=\bar{U}\Big|_{\lambda_0+\Delta\lambda} \quad\Rightarrow\quad \frac{d}{d\mu^*_{\Delta\lambda}}\bigg|_{\lambda_0+\Delta\lambda}=\frac{d}{d\mu}\bigg|_{\lambda_0+\Delta\lambda} \qquad (3.4a)$$

and is Lie dragged by $\bar{V}$ so that [see §3.3]

$$\left[\bar{U}^*_{\Delta\lambda},\bar{V}\right]=0 \quad\Rightarrow\quad \left[\frac{d}{d\mu^*_{\Delta\lambda}},\frac{d}{d\lambda}\right]=0 \qquad (3.4)$$

Now,

$$U^*_{\Delta\lambda}(f)\Big|_{\lambda_0}=\frac{df}{d\mu^*_{\Delta\lambda}}\bigg|_{\lambda_0}$$

$$=\frac{df}{d\mu^*_{\Delta\lambda}}\bigg|_{\lambda_0+\Delta\lambda}-\Delta\lambda\frac{d}{d\lambda}\left(\frac{df}{d\mu^*_{\Delta\lambda}}\right)\bigg|_{\lambda_0}+O\left(\Delta\lambda^2\right) \qquad \text{[Taylor series]}$$

$$=\frac{df}{d\mu^*_{\Delta\lambda}}\bigg|_{\lambda_0+\Delta\lambda}-\Delta\lambda\frac{d}{d\mu^*_{\Delta\lambda}}\left(\frac{df}{d\lambda}\right)\bigg|_{\lambda_0}+O\left(\Delta\lambda^2\right) \qquad \text{[(3.4) used]}$$

$$=\frac{df}{d\mu}\bigg|_{\lambda_0+\Delta\lambda}-\Delta\lambda\frac{d}{d\mu^*_{\Delta\lambda}}\left(\frac{df}{d\lambda}\right)\bigg|_{\lambda_0}+O\left(\Delta\lambda^2\right) \qquad \text{[(3.4a) used]}$$

$$=\frac{df}{d\mu}\bigg|_{\lambda_0}+\Delta\lambda\frac{d}{d\lambda}\left(\frac{df}{d\mu}\right)\bigg|_{\lambda_0}-\Delta\lambda\frac{d}{d\mu^*_{\Delta\lambda}}\left(\frac{df}{d\lambda}\right)\bigg|_{\lambda_0}+O\left(\Delta\lambda^2\right) \qquad \text{[Taylor]}$$

$$= \left.\frac{df}{d\mu}\right|_{\lambda_0} + \Delta\lambda \frac{d}{d\lambda}\left(\frac{df}{d\mu}\right)\Bigg|_{\lambda_0} - \Delta\lambda \frac{d}{d\mu}\left(\frac{df}{d\lambda}\right)\Bigg|_{\lambda_0} + O\left(\Delta\lambda^2\right)$$

where we've used $\frac{d}{d\mu^*_{\Delta\lambda}} = \frac{d}{d\mu} + O\left(\Delta\lambda\right)$ to get the last equality. Eq(3.5) thus becomes

$$\left[\mathcal{L}_{\bar{V}}\bar{U}\right](f) = \frac{d}{d\lambda}\left(\frac{df}{d\mu}\right) - \frac{d}{d\mu}\left(\frac{df}{d\lambda}\right) = \left[\frac{d}{d\lambda}, \frac{d}{d\mu}\right] f$$

Since this applies to all *f*, we have

$$\mathcal{L}_{\bar{V}}\bar{U} = \left[\frac{d}{d\lambda}, \frac{d}{d\mu}\right] = \left[\bar{V}, \bar{U}\right] \qquad (3.6)$$

Note that

$$\mathcal{L}_{\bar{U}}\bar{V} = \left[\bar{U}, \bar{V}\right] = -\mathcal{L}_{\bar{V}}\bar{U} \qquad (3.7)$$

3.4.3. Properties

The proof for the following properties of the Lie derivative are left as an exercise.

1. $\left[\mathcal{L}_{\overline{V}}, \mathcal{L}_{\overline{W}}\right] = \mathcal{L}_{\left[\overline{V}, \overline{W}\right]}$ (3.8)

2. $\mathcal{L}_{\overline{V}} + \mathcal{L}_{\overline{W}} = \mathcal{L}_{\overline{V}+\overline{W}}$

3. **Jacobi identity**:

$$\left[\left[\mathcal{L}_{\overline{U}}, \mathcal{L}_{\overline{V}}\right], \mathcal{L}_{\overline{W}}\right] + \left[\left[\mathcal{L}_{\overline{V}}, \mathcal{L}_{\overline{W}}\right], \mathcal{L}_{\overline{U}}\right] + \left[\left[\mathcal{L}_{\overline{W}}, \mathcal{L}_{\overline{U}}\right], \mathcal{L}_{\overline{V}}\right] = 0 \tag{3.9}$$

4. **Leibniz rule**:

$$\mathcal{L}_{\overline{V}}\left(f\overline{U}\right) = \left(\mathcal{L}_{\overline{V}} f\right)\overline{U} + f\mathcal{L}_{\overline{V}}\overline{U} \tag{3.10}$$

5. For coordinates $\left\{x^i\right\}$,

$$\left(\mathcal{L}_{\overline{V}}\overline{U}\right)^i = V^j \frac{\partial}{\partial x^j} U^i - U^j \frac{\partial}{\partial x^j} V^i \tag{2.7}$$

so that if $\overline{V} = \frac{\partial}{\partial x^i}$, then

$$\left(\mathcal{L}_{\overline{V}}\overline{U}\right)^i = \frac{\partial U^i}{\partial x^j} \tag{3.12}$$

6. For an arbitrary basis $\left\{\overline{e}^i\right\}$,

$$\left(\mathcal{L}_{\overline{V}}\overline{U}\right)^i = V^j \overline{e}_j\left(U^i\right) - U^j \overline{e}_j\left(V^i\right) + V^j U^k \left(\mathcal{L}_{\overline{e}_j}\overline{e}_k\right)^i \tag{3.11}$$

3.5. Lie Derivative of a One-Form

Since $\tilde{\omega}(\bar{U})$ is a function, the Leibniz rule

$$\mathcal{L}_{\bar{V}}\left[\tilde{\omega}(\bar{U})\right]=\left(\mathcal{L}_{\bar{V}}\tilde{\omega}\right)\bar{U}+\tilde{\omega}\left(\mathcal{L}_{\bar{V}}\bar{U}\right) \tag{3.13}$$

serves as the definition of the Lie derivative $\mathcal{L}_{\bar{V}}\tilde{\omega}$ of the 1-form $\tilde{\omega}$ along vector $\bar{V}$.

Thus, $\mathcal{L}_{\bar{V}}\tilde{\omega}$ is a 1-form. Given coordinates $\{x^i\}$, eq(3.13) becomes

$$V^i\frac{\partial}{\partial x^i}\left(\omega_j U^j\right)=\left(\mathcal{L}_{\bar{V}}\tilde{\omega}\right)_j U^j+\omega_j\left(V^i\frac{\partial}{\partial x^i}U^j-U^i\frac{\partial}{\partial x^i}V^j\right)$$

$\Rightarrow$

$$\begin{aligned}\left(\mathcal{L}_{\bar{V}}\tilde{\omega}\right)_j U^j &= V^i\frac{\partial\omega_j}{\partial x^i}U^j+V^i\omega_j\frac{\partial U^j}{\partial x^i}-\omega_j V^i\frac{\partial U^j}{\partial x^i}+\omega_j U^i\frac{\partial V^j}{\partial x^i}\\ &= V^i\frac{\partial\omega_j}{\partial x^i}U^j+\omega_j U^i\frac{\partial V^j}{\partial x^i}\\ &= \left(V^i\frac{\partial\omega_j}{\partial x^i}+\omega_i\frac{\partial V^i}{\partial x^j}\right)U^j\end{aligned}$$

Since this holds for arbitrary U^j, we have

$$\left(\mathcal{L}_{\bar{V}}\tilde{\omega}\right)_j=V^k\frac{\partial\omega_j}{\partial x^k}+\omega_k\frac{\partial V^k}{\partial x^j} \tag{3.14}$$

which should be compared with

$$\left(\mathcal{L}_{\bar{V}}\bar{U}\right)^j=V^k\frac{\partial U^j}{\partial x^k}-U^k\frac{\partial V^j}{\partial x^k} \tag{2.7}$$

To summarize, the Lie derivative of a vector (1-form) is a vector (1-form). Since a general tensor of type $\begin{pmatrix}N\\M\end{pmatrix}$ can be written as a linear combination of the direct products of N vectors and M 1-forms, the Lie derivative preserves the type of the tensor it operates on.

For arbitrary tensors **A**, **B**, and **T**, the Leibniz rule becomes

$$\mathcal{L}_{\bar{V}}\left(\mathbf{A}\otimes\mathbf{B}\right)=\left(\mathcal{L}_{\bar{V}}\mathbf{A}\right)\otimes\mathbf{B}+\mathbf{A}\otimes\left(\mathcal{L}_{\bar{V}}\mathbf{B}\right) \tag{3.15}$$

so that (3.13) generalizes to

$$\mathcal{L}_{\bar{V}}\left[\mathbf{T}\left(\tilde{\omega},\cdots,\bar{U},\cdots\right)\right]=\left(\mathcal{L}_{\bar{V}}\mathbf{T}\right)\left(\tilde{\omega},\cdots,\bar{U},\cdots\right)$$

$$+\mathbf{T}\left(\mathcal{L}_{\overline{V}}\tilde{\omega},\cdots,\overline{U},\cdots\right)+\cdots$$

$$+\mathbf{T}\left(\tilde{\omega},\cdots,\mathcal{L}_{\overline{V}}\overline{U},\cdots\right)+\cdots \tag{3.16}$$

3.6. Submanifolds

Loosely speaking, a submanifold is a subset of a manifold that is itself a manifold. Sometimes it is called a hypersurface. Indeed, many non-equivalent definitions exists in the literature so that care must be exercised in using the term.

Here, we define an *m*-D **submanifold** *S* of an *n*-D manifold *M* as a smooth subset of *M* such that for every region $U \subset S \subset M$, it is possible to find a coordinate patch in *M* such that the coordinates of every point in *U* has the form $\left(x^1,\cdots,x^m,a^1,\cdots,a^{n-m}\right)$, where a^j are constants that may for convenience set to 0.

Solutions to differential equations that are in the form

$$\left\{y_i = y_i\left(x^1,\cdots,x^m\right);\, i=1,\cdots,p\right\}$$

can be thought of as *m*-D submanifolds with coordinates $\left(x^1,\cdots,x^m\right)$ in an (*m*+*p*)-D manifold with coordinates $\left\{x^1,\cdots,x^m,y_1,\cdots,y_p\right\}$. [see chapter 4]

Obviously, a curve in *S* through a point *P* is also a curve in *M*. Every vector in the tangent space V_P of *S* is therefore also a vector in the tangent space T_P of *M*. Hence, V_P is a *m*-D subspace of T_P. Note that a vector in T_P has no unique projection on V_P.

The situation for 1-forms at *P* is just the reverse. Thus, any 1-form in cotangent space T_P^* of *M* is a 1-form in V_P^* of *S* so that T_P^* is a subspace of V_P^*. A 1-form in V_P^* has no unique projection on T_P^*.

3.7. Frobenius' Theorem (Vector Field Version)

In any coordinate patch of a submanifold *S* with coordinates $\{y^a \, , a = 1, \cdots, m\}$, there are basis vectors $\left\{ \frac{\partial}{\partial y^a} \right\}$ for vector fields on *S* with

$$\left[\frac{\partial}{\partial y^a} , \frac{\partial}{\partial y^b} \right] = 0 \qquad (3.17)$$

It is left as an exercise to show that

a) If $\bar{V}$ and $\bar{W}$ are linear combinations of *m* commuting vector fields $\{\bar{U}_{(j)} \, , \, j = 1, \cdots, m\}$ using functions as coefficients, then $[\bar{V}, \bar{W}]$ is a linear combination of $\{\bar{U}_{(j)}\}$. The set $\{\bar{U}_{(j)}\}$ is said to be closed with respect to the Lie bracket.

b) The same holds when $\{\bar{U}_{(j)}\}$ have Lie brackets that are non-vanishing linear combinations of $\{\bar{U}_{(j)}\}$.

It follows from (a) that the Lie bracket of any 2 vector fields on *S* is also tangent to *S*. Conversely, the **Frobenius's theorem** says that

> If a set of *m* C^∞ vector fields defined in a region *U* of *M* is closed with respect to the Lie bracket, the integral curves of the fields mesh to form a family of submanifolds.

The family of submanifolds fill up *U* the same way as a congruence of curves does. It is called a **foliation** of *U* and each manifold is called a **leaf** of the foliation. [see figs 3.6,7]

There is another version of Frobenius's theorem stated in terms of differential forms [see §4.26]. It is the fundamental theorem for obtaining integrability conditions for partial differential equations.

3.8. Proof of Frobenius' Theorem

3.8.1. Strategy of Proof

Let M be a manifold of dimension n. Let $\{\bar{V}_{(i)}\ ; i=1,\cdots,m'\}$ be a set of m' vector fields defined in a region $U \subset M$. Closure under the Lie bracket means:

$$\left[\bar{V}_{(i)},\bar{V}_{(j)}\right]=\sum_k \alpha_{ijk}\bar{V}_{(k)} \qquad \forall\ i,j,k \in (1,\cdots,m')$$

where the parenthesis around the indices are used to distinguish them from tensor indices. If $\{\bar{V}_{(i)}\}$ spans a subspace of $T_P(U)$ of dimension $m \le m' < n$, only m of the vectors are linearly independent. Re-labeling them as $\{\bar{V}_{(a)}\ ; a=1,\cdots,m\}$, we have

$$\left[\bar{V}_{(a)},\bar{V}_{(b)}\right]=\sum_c \alpha_{abc}\bar{V}_{(c)} \qquad \forall\ a,b,c \in (1,\cdots,m)$$

For each vector $\bar{V}$ in $T_P(M)$, there is a curve in M whose tangent at P is $\bar{V}$. For each vector field $\bar{V}$ on U, there is an integral curve whose tangent at every point is $\bar{V}$. Any two independent vectors $\bar{V}$ and $\bar{W}$ will form a 2-D subspace of $T_P(M)$. Likewise, two independent vector fields $\bar{V}$ and $\bar{W}$ will form a 2-D subspace of $T(M)$. However, there is no reason why the corresponding integral curves also lie in a 2-D subspace (submanifold) of M. What the Frobenius theorem means is that if $\{\bar{V}_{(a)}\}$ is closed under the Lie bracket, their corresponding integral curves will lie entirely in a submanifold of M. This then implies M is sub-divided into a family of such non-intersecting submanifolds, i.e., a **foliation**. If

$$\left[\bar{V}_{(a)},\bar{V}_{(b)}\right]=0 \qquad \forall\ a,b$$

the set $\{\bar{V}_{(a)}\ ; a=1,\cdots,m\}$ forms a coordinate basis for a submanifold S of U. The theorem is then trivially proved. For the case

$$\left[\bar{V}_{(a)},\bar{V}_{(b)}\right]\neq 0$$

we need only show that a set of commuting vector fields $\{\bar{X}_{(a)}\}$ can be constructed by the linear combinations of $\{\bar{V}_{(a)}\}$ & the proof will be completed.

3.8.2. Preliminary Relations

3.8.2.a. Closure Theorem

The Lie bracket of linear combinations of a set of vectors that are closed under the Lie bracket is still closed. (cf. Ex 3.5, Schutz) To be more explicit, let

$\left\{ \bar{U}_{(i)}, i = 1,\cdots,m \right\}$ be closed under the Lie bracket, i.e.,

$$\left[\bar{U}_{(i)}, \bar{U}_{(j)} \right] = C_{ij}^{k} \bar{U}_{(k)}$$

Let $\bar{V}$ and $\bar{W}$ be arbitrary linear combinations of $\left\{ \bar{U}_{(i)} \right\}$, i.e.,

$$\bar{V} = a^{i} \bar{U}_{(i)} \qquad \text{and} \qquad \bar{W} = b^{i} \bar{U}_{(i)}$$

where *a*, *b* can be functions, then

$$\left[\bar{V}, \bar{W} \right] = c^{i} \bar{U}_{(i)}$$

Proof:

Since *a*, *b* are functions and the vectors are derivations, we have

$$\begin{aligned}
\left[\bar{V}, \bar{W} \right] &= a^{i} \bar{U}_{(i)} b^{j} \bar{U}_{(j)} - b^{j} \bar{U}_{(j)} a^{i} \bar{U}_{(i)} \\
&= a^{i} \left(\bar{U}_{(i)} b^{j} \right) \bar{U}_{(j)} + a^{i} b^{j} \bar{U}_{(i)} \bar{U}_{(j)} - b^{j} \left(\bar{U}_{(j)} a^{i} \right) \bar{U}_{(i)} - b^{j} a^{i} \bar{U}_{(j)} \bar{U}_{(i)} \\
&= \left\{ a^{i} \left(\bar{U}_{(i)} b^{j} \right) - b^{i} \left(\bar{U}_{(i)} a^{j} \right) \right\} \bar{U}_{(j)} + a^{i} b^{j} \left[\bar{U}_{(i)}, \bar{U}_{(j)} \right] \\
&= \left\{ a^{i} \left(\bar{U}_{(i)} b^{j} \right) - b^{i} \left(\bar{U}_{(i)} a^{j} \right) \right\} \bar{U}_{(j)} + a^{i} b^{j} C_{ij}^{k} \bar{U}_{(k)} \\
&= \left\{ a^{i} \left(\bar{U}_{(i)} b^{k} \right) - b^{i} \left(\bar{U}_{(i)} a^{k} \right) + a^{i} b^{j} C_{ij}^{k} \right\} \bar{U}_{(k)} \\
&\equiv c^{k} \bar{U}_{(k)}
\end{aligned}$$

where

$$c^{k} = a^{i} \left(\bar{U}_{(i)} b^{k} \right) - b^{i} \left(\bar{U}_{(i)} a^{k} \right) + a^{i} b^{j} C_{ij}^{k}$$

3.8.2.b. Commutativity

The Lie derivative commutes with the exterior derivative, i.e.,

$$\mathcal{L}_{\bar{V}}\left(\tilde{d}\, f\right)=\tilde{d}\left(\mathcal{L}_{\bar{V}}\, f\right)$$

Proof:

Given a coordinate basis $\{\bar{e}_i\}$, we have

$$\bar{V}=V^i\bar{e}_i \qquad \tilde{d}\, f=\frac{\partial f}{\partial x^i}\tilde{e}^i$$

$$\begin{aligned}\mathcal{L}_{\bar{V}}\left(\tilde{d}\, f\right)&=V^i\mathcal{L}_{\bar{e}_i}\left(\frac{\partial f}{\partial x^j}\tilde{e}^j\right)\\&=V^i\left(\mathcal{L}_{\bar{e}_i}\frac{\partial f}{\partial x^j}\right)\tilde{e}^j+V^i\frac{\partial f}{\partial x^j}\mathcal{L}_{\bar{e}_i}\tilde{e}^j\\&=V^i\frac{\partial^2 f}{\partial x^i\partial x^j}\tilde{e}^j+\frac{\partial f}{\partial x^j}\mathcal{L}_{\bar{V}}\tilde{e}^j \qquad \text{(a)}\end{aligned}$$

$$\tilde{d}\left(\mathcal{L}_{\bar{V}}\, f\right)=\tilde{d}\left(V^i\frac{\partial f}{\partial x^i}\right)=\left(\frac{\partial V^i}{\partial x^j}\frac{\partial f}{\partial x^i}+V^i\frac{\partial^2 f}{\partial x^j\partial x^i}\right)\tilde{e}^j \qquad \text{(b)}$$

Now,

$$\left(\mathcal{L}_{\bar{V}}\tilde{e}^j\right)_i=V^k\frac{\partial}{\partial x^k}\delta_i^j+\delta_k^j\frac{\partial}{\partial x^i}V^k=\frac{\partial V^j}{\partial x^i}$$

$$\mathcal{L}_{\bar{V}}\tilde{e}^j=\frac{\partial V^j}{\partial x^i}\tilde{e}^i$$

so that (a) = (b), which completes the proof.

3.8.2.c. Inner Product

$$\left\langle \tilde{d}f \;, \left[\bar{V} \;, \bar{W}\right]\right\rangle = \mathcal{L}_{\bar{V}} \left\langle \tilde{d}f \;, \bar{W}\right\rangle - \left\langle \tilde{d}\left(\mathcal{L}_{\bar{V}} f\right), \bar{W}\right\rangle$$

Proof:

Leibniz's rule gives

$$\mathcal{L}_{\bar{V}} \left\langle \tilde{d}f \;, \bar{W}\right\rangle = \left\langle \mathcal{L}_{\bar{V}}\tilde{d}f \;, \bar{W}\right\rangle + \left\langle \tilde{d}f \;, \mathcal{L}_{\bar{V}}\bar{W}\right\rangle$$

$$= \left\langle \tilde{d}\mathcal{L}_{\bar{V}} f \;, \bar{W}\right\rangle + \left\langle \tilde{d}f \;, \left[\bar{V} \;, \bar{W}\right]\right\rangle$$

QED.

3.8.3. Proof

The proof is by induction.
The case $m = 1$ is just the congruence with each leaf an integral curve.
Next, we show that case $m-1$ is true implies case *m* is true.

Out of the linearly independent fields $\left\{ \overline{V}_{(a)} \; ; a = 1, \cdots, m \right\}$ with

$$\left[\overline{V}_{(a)} \, , \overline{V}_{(b)} \right] = \sum_c \alpha_{abc} \overline{V}_{(c)} \qquad \forall \; a, b, c \in (1, \cdots, m)$$

we choose one and call it $\overline{V}_{(m)} = \dfrac{d}{d\lambda_{(m)}}$. We then construct $\left\{ \overline{X}_{(a)} \; ; a = 1, \cdots, m-1 \right\}$

using linear combinations of $\left\{ \overline{V}_{(a)} \right\}$ so that

$$\left\langle \tilde{d}\lambda_{(m)} \, , \overline{X}_{(a)} \right\rangle = 0 \qquad \forall \; a = 1, \cdots, m-1$$

This is always possible since $\tilde{d}\lambda_{(m)}$ is a 1-form of the entire *n*-D manifold *M* so that it can be "orthogonal" up to *n* – 1 independent vector fields. Note that $n \geq m' \geq m$.

The set of *m* fields $\left\{ \overline{V}_{(m)}, \overline{X}_{(a)} \; ; a = 1, \cdots, m-1 \right\}$ is linearly independent with coefficients that can be functions. According to the closure theorem, this set is still closed under the Lie bracket so that we can write

$$\left[\overline{X}_{(a)} \, , \overline{X}_{(b)} \right] = \sum_{c=1}^{m-1} \beta_{abc} \overline{X}_{(c)} + \gamma_{ab} \overline{V}_{(m)}$$

$$\left[\overline{V}_{(m)} \, , \overline{X}_{(b)} \right] = \sum_{b=1}^{m-1} \mu_{ab} \overline{X}_{(b)} + \nu_a \overline{V}_{(m)}$$

Now,

$$\left\langle \tilde{d}\lambda_{(m)} \, , \overline{V}_{(m)} \right\rangle = \tilde{d}\lambda_{(m)} \left(\frac{d}{d\lambda_{(m)}} \right) = \frac{d\lambda_{(m)}}{d\lambda_{(m)}} = 1$$

$$\mathcal{L}_{\overline{V}_{(m)}} \lambda_{(m)} = \frac{d\lambda_{(m)}}{d\lambda_{(m)}} = 1$$

⤢ $$\tilde{d} \left\langle \tilde{d}\lambda_{(m)} \, , \overline{V}_{(m)} \right\rangle = \tilde{d} \left(\mathcal{L}_{\overline{V}_{(m)}} \lambda_{(m)} \right) = \tilde{d} \, 1 = 0$$

Thus:

$$\left\langle \tilde{d}\lambda_{(m)} , \left[\overline{X}_{(a)} , \overline{X}_{(b)} \right] \right\rangle = \mathcal{L}_{\overline{X}_{(a)}} \left\langle \tilde{d}\lambda_{(m)} , \overline{X}_{(b)} \right\rangle - \left\langle \tilde{d}\left(\mathcal{L}_{\overline{X}_{(a)}} \lambda_{(m)} \right), \overline{X}_{(b)} \right\rangle$$

$$= -\left\langle \tilde{d}\left(\mathcal{L}_{\overline{X}_{(a)}} \lambda_{(m)} \right), \overline{X}_{(b)} \right\rangle = 0$$

where $\overline{X}_{(a)} = \dfrac{d}{d\mu_a}$ and $\mathcal{L}_{\overline{X}_{(a)}} \lambda_{(m)} = \dfrac{d\lambda_{(m)}}{d\mu_a} = 0$ since $\overline{V}_{(m)}$ and $\overline{X}_{(a)}$ are independent.

Similarly,

$$\left\langle \tilde{d}\lambda_{(m)} , \left[\overline{V}_{(m)} , \overline{X}_{(b)} \right] \right\rangle = \mathcal{L}_{\overline{V}_{(m)}} \left\langle \tilde{d}\lambda_{(m)} , \overline{X}_{(b)} \right\rangle - \left\langle \tilde{d}\left(\mathcal{L}_{\overline{V}_{(m)}} \lambda_{(m)} \right), \overline{X}_{(b)} \right\rangle$$

$$= -\left\langle \tilde{d}\left(\mathcal{L}_{\overline{V}_{(m)}} \lambda_{(m)} \right), \overline{X}_{(b)} \right\rangle = 0$$

Using $\left\langle \tilde{d}\lambda_{(m)} , \overline{X}_{(a)} \right\rangle = 0$, contraction of the closure conditions with $\tilde{d}\lambda_{(m)}$ becomes

$$0 = \gamma_{ab} \left\langle \tilde{d}\lambda_{(m)} , \overline{V}_{(m)} \right\rangle$$

$$0 = \nu_a \left\langle \tilde{d}\lambda_{(m)} , \overline{V}_{(m)} \right\rangle$$

Hence:

$\gamma_{ab} = 0$ and $\nu_a = 0$

so that

$$\left[\overline{X}_{(a)} , \overline{X}_{(b)} \right] = \sum_{c=1}^{m-1} \beta_{abc} \overline{X}_{(c)}$$

$$\left[\overline{V}_{(m)} , \overline{X}_{(b)} \right] = \sum_{b=1}^{m-1} \mu_{ab} \overline{X}_{(b)}$$

Thus, $\left\{ \overline{X}_{(a)} \,;\, a = 1, \cdots, m-1 \right\}$ alone is closed under the Lie bracket, which means they form a foliation of the (m–1)-D submanifold. (Remember that we assumed case m–1 of the theorem to be true) By definition, a foliation always has a coordinate basis. Thus, a linear combination $\left\{ \overline{Y}_{(a)} \right\}$ of $\left\{ \overline{X}_{(a)} \right\}$ exists such that

$$\left[\overline{Y}_{(a)} , \overline{Y}_{(b)} \right] = 0 \qquad\qquad \forall\, a, b = 1, \cdots, m-1$$

Now, if we can construct $\left\{ \overline{Z}_{(a)} \,;\, a = 1, \cdots, m-1 \right\}$ from $\left\{ \overline{X}_{(a)} \right\}$ such that

$$\overline{Z}_{(a)} = \sum_{b=1}^{m-1} \alpha_{ab} \overline{X}_{(b)}$$

$$\left[\overline{V}_{(m)} , \overline{Z}_{(b)} \right] = 0$$

$$\left[\bar{Z}_{(a)}, \bar{Z}_{(b)}\right]=0 \qquad \forall\, a,b=1,\cdots,m-1$$

the proof will be completed. ($\left\{\bar{V}_{(m)},\bar{Z}_{(a)}\right\}$ will be the coordinate basis of the *m*-D submanifold).

Starting with $\left[\bar{V}_{(m)},\bar{Z}_{(a)}\right]=0$, we have

$$0=\left[\bar{V}_{(m)}, \bar{Z}_{(a)}\right] = \mathcal{L}_{\bar{V}_{(m)}}\bar{Z}_{(a)} = \mathcal{L}_{\bar{V}_{(m)}}\sum_{b=1}^{m-1}\alpha_{ab}\bar{X}_{(b)}$$

$$=\sum_{b=1}^{m-1}\left(\mathcal{L}_{\bar{V}_{(m)}}\alpha_{ab}\right)\bar{X}_{(b)}+\sum_{b=1}^{m-1}\alpha_{ab}\mathcal{L}_{\bar{V}_{(m)}}\bar{X}_{(b)}$$

$$=\sum_{b=1}^{m-1}\frac{d\alpha_{ab}}{d\lambda_{(m)}}\bar{X}_{(b)}+\sum_{b=1}^{m-1}\alpha_{ab}\left[\bar{V}_{(m)}, \bar{X}_{(b)}\right]$$

$$=\sum_{b=1}^{m-1}\frac{d\alpha_{ab}}{d\lambda_{(m)}}\bar{X}_{(b)}+\sum_{b=1}^{m-1}\alpha_{ab}\sum_{c=1}^{m-1}\mu_{bc}\bar{X}_{(c)}$$

$$=\sum_{b=1}^{m-1}\left(\frac{d\alpha_{ab}}{d\lambda_{(m)}}+\sum_{c=1}^{m-1}\alpha_{ac}\mu_{cb}\right)\bar{X}_{(b)}$$

Since $\bar{X}_{(b)}$ are independent, we have

$$\frac{d\alpha_{ab}}{d\lambda_{(m)}}+\sum_{c=1}^{m-1}\alpha_{ac}\mu_{cb}=0$$

which is a set of ordinary differential eqs for α. (μ is fixed for given $\bar{V}_{(m)}$ and $\left\{\bar{X}_{(a)}\right\}$). Since a solution for α always exists for a given set of initial conditions, the condition $\left[\bar{V}_{(m)},\bar{Z}_{(a)}\right]=0$ can always be satisfied. The initial conditions are chosen such that $\left\{\bar{Z}_{(a)}\right\}$ are coordinate basis of the (*m*–1)-D submanifold spanned by $\left\{\bar{X}_{(a)}\right\}$ so that

$$\left[\bar{Z}_{(a)}, \bar{Z}_{(b)}\right]=0$$

on the submanifold. This is accomplished by setting $\bar{Z}_{(a)}=\bar{Y}_{(a)}$. Outside the submanifold, the condition

$$0 = \left[\bar{V}_{(m)}, \bar{Z}_{(a)}\right] = \mathcal{L}_{\bar{V}_{(m)}} \bar{Z}_{(a)}$$

means that $\bar{Z}_{(a)}$ is the Lie dragging of $\bar{Y}_{(a)}$ along $\bar{V}_{(m)}$. Since $\left[\bar{Y}_{(a)}, \bar{Y}_{(b)}\right] = 0$ we have $\left[\bar{Z}_{(a)}, \bar{Z}_{(b)}\right] = 0$ because Lie dragging preserves the commutativity of Lie brackets. QED.

3.9. An Example: The Generators of S^2

Consider the (un-normalized) basis vector $\overline{e}_\phi$ of the spherical coordinates (r,θ,ϕ),

$$\overline{e}_\phi = -y\overline{e}_x + x\overline{e}_y \qquad \text{or} \qquad \frac{\partial}{\partial\phi} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$$

which we shall called $\overline{l}_z$, the **angular momentum operator** along *z*,

$$\overline{l}_z = \frac{\partial}{\partial\phi}$$

Defining $\overline{l}_x$ and $\overline{l}_y$ analogously, one can prove that [see any quantum mechanics textbook],

$$\left[\overline{l}_x , \overline{l}_y\right] = -\overline{l}_z$$

$$\left[\overline{l}_y , \overline{l}_z\right] = -\overline{l}_x \qquad (3.30)$$

$$\left[\overline{l}_z , \overline{l}_x\right] = -\overline{l}_y$$

Thus, $\left\{\overline{l}_x, \overline{l}_y, \overline{l}_z\right\}$ is closed under the Lie bracket so that they form a submanifold of dimension $d \le 3$. In fact, $d = 2$, which can be seen by considering the function

$$r = \sqrt{x^2 + y^2 + z^2}$$

It can be shown that

$$\overline{l}_x(r) = \overline{l}_y(r) = \overline{l}_z(r) = 0$$

and

$$\tilde{d}r\left(\overline{l}_x\right) = \tilde{d}r\left(\overline{l}_y\right) = \tilde{d}r\left(\overline{l}_z\right) = 0 \qquad (3.31)$$

Since the contraction $\tilde{\omega}\left(\overline{V}\right)$ is the number of planes of constant ω that are pierced by $\overline{V}$, eq(3.31) means that $\left\{\overline{l}_x, \overline{l}_y, \overline{l}_z\right\}$ are tangent to the 2-D surface $r = const$. Therefore, $d = 2$ and the submanifold is just the 2-D sphere.

To complete the analogy with the angular momentum operator in quantum mechanics, we define

$$L^2 = \mathcal{L}_{\overline{l_x}}\mathcal{L}_{\overline{l_x}} + \mathcal{L}_{\overline{l_y}}\mathcal{L}_{\overline{l_y}} + \mathcal{L}_{\overline{l_z}}\mathcal{L}_{\overline{l_z}} \tag{3.32}$$

so that

$$\left[\mathcal{L}_{\overline{l_j}}, L^2\right] = 0 \qquad \text{for} \quad j = x, y, z$$

$$L^2 f = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2} \tag{3.33}$$

Proof of these is left as exercise.

3.10. Invariance

A tensor field **T** is said to be **invariant under a vector field** $\overline{V}$ if

$$\mathcal{L}_{\overline{V}}\mathbf{T} = 0 \qquad (3.34)$$

If **T** is an important characteristic of the system, so is $\overline{V}$. For example, an **axially symmetric** system is invariant under rotations in some plane so that the angular momentum generating them is conserved and characterizes the system.

Theorem

Let $F = \{\mathbf{T}_{(1)}, \mathbf{T}_{(2)}, \cdots\}$ be a set of tensor fields. The set *L* of all vector fields under which all fields in *F* are invariant is a Lie algebra.

Proof

First, we leave as an exercise to show that

$$\mathcal{L}_{\overline{V}}\mathbf{T}_{(i)} = \mathcal{L}_{\overline{W}}\mathbf{T}_{(i)} = 0 \quad \Rightarrow \quad \mathcal{L}_{a\overline{V}+b\overline{W}}\mathbf{T}_{(i)} = 0 \qquad (3.35a)$$

where *a* and *b* are constants. Secondly, using eq(3.8), we have

$$\mathcal{L}_{\overline{V}}\mathbf{T}_{(i)} = \mathcal{L}_{\overline{W}}\mathbf{T}_{(i)} = 0 \quad \Rightarrow \quad \left[\mathcal{L}_{\overline{V}}, \mathcal{L}_{\overline{W}}\right]\mathbf{T}_{(i)} = \mathcal{L}_{[\overline{V},\overline{W}]}\mathbf{T}_{(i)} = 0 \qquad (3.35)$$

Thus, if $\overline{V}$ and $\overline{W}$ are in *L*, so are $a\overline{V} + b\overline{W}$ and $\left[\overline{V},\overline{W}\right]$. QED.

Comments

Note that *L* allows only linear combinations with *constant* coefficients. Thus, the corresponding vector space treats each *field* $\overline{V}$ as a single element. In contrast, a fibre bundle allows linear combinations with *functions* as coefficients so that each field $\overline{V}$ is a cross section. For example, the set $\{\overline{l}_x, \overline{l}_y, \overline{l}_z\}$ as considered vectors in the fibre bundle with base $\mathbb{R}^3$ is linearly dependent since they are all tangent to the spherical surface S^2. However, to express each vector field $\overline{l}_j$ as the linear combination of the other 2, we must use functions as coefficients. Hence, in the Lie algebra, the set $\{\overline{l}_x, \overline{l}_y, \overline{l}_z\}$ is linearly independent and form a 3-D basis.

Similarly, the dimension of the Lie algebra of all tangent vector fields to a finite dimensional manifold is infinite.

3.11. Killing Vector Fields

A **Killing vector field** $\bar{V}$ is defined by

$$\mathcal{L}_{\bar{V}}\, \mathfrak{g} = 0 \tag{3.36}$$

where $\mathfrak{g}$ is the metric tensor. Given coordinates $\{x^i\}$, we have [see eq(3.14)]

$$\left(\mathcal{L}_{\bar{V}}\mathfrak{g}\right)_{ij} = V^k \frac{\partial}{\partial x^k} g_{ij} + g_{ik} \frac{\partial}{\partial x^j} V^k + g_{kj} \frac{\partial}{\partial x^j} V^k = 0 \tag{3.37}$$

If the integral curves of $\bar{V}$ are used as the coordinate lines of x^1, eq(3.37) reduces to [see eq(3.12)]

$$\left(\mathcal{L}_{\bar{V}}\mathfrak{g}\right)_{ij} = \frac{\partial}{\partial x^1} g_{ij} = 0 \tag{3.38}$$

Thus, if a Killing vector is used as a basis vector, the metric will be independent of the corresponding coordinate. Conversely, if a metric is independent of certain coordinates, the corresponding basis vectors are Killing vectors.

As an example, consider $\mathbb{E}^3$ with Cartesian coordinates $\{x, y, z\}$ so that

$$g_{ij} = \delta_{ij} \quad \text{for} \quad i, j = x, y, z \tag{3.39}$$

Since these are independent of $\{x, y, z\}$, the basis vectors $\left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right\}$ are Killing vectors. In spherical coordinates $\{r, \theta, \phi\}$, we have

$$\begin{aligned} g_{rr} &= \frac{\partial}{\partial r} \cdot \frac{\partial}{\partial r} = 1 \\ g_{\theta\theta} &= \frac{\partial}{\partial \theta} \cdot \frac{\partial}{\partial \theta} = r^2 \\ g_{\phi\phi} &= \frac{\partial}{\partial \phi} \cdot \frac{\partial}{\partial \phi} = r^2 \sin^2\theta \end{aligned} \tag{3.40}$$

Hence, only $\frac{\partial}{\partial \phi} = \bar{l}_z$ is a Killing vector. By symmetry, so are $\bar{l}_x$ and $\bar{l}_y$. Indeed, the 6 Killing vectors $\{x, y, z, \bar{l}_x, \bar{l}_y, \bar{l}_z\}$ form a basis for the Lie algebra (of Killing vector fields). [More in chapter 5]

3.12. Killing Vectors and Conserved Quantities in Particle Dynamics

In classical mechanics, if a particle is subject to an axially symmetric potential, the component of its angular momentum about the axis of symmetry is a constant on the particle's trajectory. Similarly, if the potential is independent of the Cartesian coordinate *x*, the *x*-component of the linear momentum is conserved.

However, not every symmetry of the potential can lead to a conserved dynamical quantity. For example, there is no conserved quantity associated with a potential that is constant on a family of ellipsoidal surfaces.

The reason for this is that, in order to induce a conserved dynamical quantity, the symmetry of the potential must be with respect to displacements along some Killing vector fields of the spacetime manifold. Proof of this will be deferred to §5.8.

In the meantime, consider the equation of motion in ordinary vector calculus notations,

$$m\dot{\mathbf{V}} = -\nabla\Phi \qquad \text{or} \qquad m\dot{V}^i = -\nabla^i\Phi \tag{3.41}$$

where ∇ is the "gradient" operator. However, as shown in §2.29, ∇ is really the vector gradient so that in non-Cartesian coordinates, eq(3.41) should be written as

$$m\dot{V}^i = -g^{ij}\frac{\partial}{\partial x^j}\Phi \tag{3.42}$$

Obviously, any conserved quantities of (3.42) must involve symmetries in both $\mathfrak{g}$ and Φ.

3.13. Axial Symmetry

3.13.0. The Problem

Axial symmetry denotes invariance under arbitrary rotations about a fixed axis. If there is an additional invariance under arbitrary translations along the axis, the symmetry becomes **cylindrical**.

In problems such as particle dynamics, the symmetry is in the "background". For example, in the motion of a particle subject to a potential with axial symmetry, the equation of motion is of the form

$$L(\psi) = 0 \qquad (3.43)$$

where ψ describes the state of the system. L is some *linear* operator with axial symmetry, i.e., it is invariant under transformation $\phi \to \phi + const$, where ϕ is the angle about the symmetry axis. In other words, L can only depend on partials such as $\frac{\partial^n}{\partial \phi^n}$ but not ϕ itself. However, a solution ψ to (3.43) cannot be axially symmetric since the particle can only be at one place at any given time.

Similar conclusions can also be drawn for the case of adding to an axially symmetric field a perturbation that has non-axisymmetric initial values.

3.13.1. Scalar Solutions

If ψ is a scalar, one can introduce the Fourier series

$$\psi\left(\phi, x^j\right) = \sum_{m=-\infty}^{\infty} \psi_m\left(x^j\right) e^{im\phi} \qquad (3.44)$$

so that (3.43) becomes

$$\begin{aligned} L(\psi) &= \sum_{m=-\infty}^{\infty}\left[L(\psi_m) e^{im\phi} + \psi_m L\left(e^{im\phi}\right)\right] \\ &= \sum_{m=-\infty}^{\infty}\left[L(\psi_m) + e^{-im\phi}\psi_m L\left(e^{im\phi}\right)\right] e^{im\phi} \\ &= 0 \qquad (a) \end{aligned}$$

Since L is a linear differential operator, we have

$$L\left(e^{im\phi}\right) = f_m\left(x^j\right) e^{im\phi}$$

where f_m is some function independent of ϕ. Hence, (a) becomes

$$0 = \sum_{m=-\infty}^{\infty}\left[L(\psi_m) + \psi_m f_m\right] e^{im\phi} \qquad (b)$$

Since the terms enclosed in [ ] are independent of ϕ, eq(b) can only be satisfied if

$$L(\psi_m) + \psi_m f_m = L(\psi_m) + \psi_m e^{-im\phi} L\left(e^{im\phi}\right) = 0 \qquad \forall\, m$$

Thus, we can define a ϕ-independent operator L_m by

$$L_m(\psi_m) = L(\psi_m) + \psi_m e^{-im\phi} L\left(e^{im\phi}\right) = 0$$

so that

$$L(\psi) = \sum_{m=-\infty}^{\infty} e^{im\phi} L_m(\psi_m) = 0$$

Another way to write $L_m \psi_m$ is

$$\begin{aligned} L_m(\psi_m) &= e^{-im\phi}\left[e^{im\phi} L(\psi_m) + \psi_m L\left(e^{im\phi}\right)\right] \\ &= e^{-im\phi} L\left(\psi_m e^{im\phi}\right) \qquad (3.45) \end{aligned}$$

Note that the difference between L_m and L is that the former is completely independent of ϕ but the latter can depend on the partials of ϕ. For example, consider the Laplacian operator in ordinary vector calculus written in spherical coordinates

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial}{\partial\theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial\phi^2}$$

Obviously, it is invariant under $\phi \to \phi + const$. Also,

$$\nabla^2\left[f(r,\theta)e^{im\phi}\right]=\nabla\cdot\left[(\nabla f)e^{im\phi}+f\nabla e^{im\phi}\right]$$

$$=\left(\nabla^2 f\right)e^{im\phi}+2(\nabla f)\cdot\nabla e^{im\phi}+f\nabla^2 e^{im\phi}$$

$$=\left(\nabla^2 f\right)e^{im\phi}+f\nabla^2 e^{im\phi}$$

$$=\left(\nabla^2 f\right)e^{im\phi}-f\frac{m^2}{r^2\sin^2\theta}e^{im\phi}$$

$$=e^{im\phi}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)-\frac{m^2}{r^2\sin^2\theta}\right]f \qquad (3.46)$$

Thus, $L=\nabla^2$ and

$$L_m=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)-\frac{m^2}{r^2\sin^2\theta}$$

The functions $e^{im\phi}$ are called **scalar axial harmonics**. A solution ψ of $L(\psi)=0$ is said to have **axial eigenvalue** *m* if

$$\mathcal{L}_{\overline{e}_\phi}\psi=im\psi \qquad \text{where} \qquad \overline{e}_\phi=\frac{\partial}{\partial\phi}=\overline{l}_z \qquad (3.47)$$

If ψ is a scalar function,

$$\mathcal{L}_{\overline{e}_\phi}\psi=\frac{\partial\psi}{\partial\phi}=im\psi$$

so that ψ is simply proportional to the scalar axial harmonics.

3.13.2. Vector Solutions

Consider the submanifold *S* defined by $\phi = 0$ with the symmetry axis as a boundary. Let $\{\overline{e}_j\}$ be the basis for the tangent space V_P of *S*. It is then supplemented so that $\{\overline{e}_\phi, \overline{e}_j\}$ is the basis of the tangent space T_P of the manifold *M* for points in *S*. A basis of T_P for the entire *M* can be obtained by Lie dragging $\{\overline{e}_\phi, \overline{e}_j\}$ along $\overline{e}_\phi$ once around the symmetry axis. [see Fig.3.8] By definition, the resulting basis vectors satisfy

$$\mathcal{L}_{\overline{e}_\phi} \overline{e}_\phi = 0 \qquad \mathcal{L}_{\overline{e}_\phi} \overline{e}_j = 0 \qquad (3.48)$$

i.e., they are all axially symmetric. Note that the Cartesian coordinate components of these basis vectors are changed by the Lie dragging. This nicely illustrates the fact that axial symmetry for a vector field demands ϕ- independence of its components only when ϕ is one of the coordinates.

The basis generated by (3.48) has axial eigenvalue $m = 0$. Basis vectors $\{\overline{e}_{(m)\phi}, \overline{e}_{(m)j}\}$ with axial eigenvalue *m* can be obtained from $\{\overline{e}_{(0)\phi}, \overline{e}_{(0)j}\} = \{\overline{e}_\phi, \overline{e}_j\}$ according to

$$\overline{e}_{(m)\phi} = e^{im\phi}\, \overline{e}_\phi \qquad \text{and} \qquad \overline{e}_{(m)j} = e^{im\phi}\, \overline{e}_j \qquad (3.49)$$

so that using (3.48), we have

$$\mathcal{L}_{\overline{e}_\phi} \overline{e}_{(m)\phi} = \mathcal{L}_{\overline{e}_\phi}\left(e^{im\phi}\, \overline{e}_\phi\right) = \left(\mathcal{L}_{\overline{e}_\phi} e^{im\phi}\right)\overline{e}_\phi + e^{im\phi}\mathcal{L}_{\overline{e}_\phi}\left(\overline{e}_\phi\right)$$

$$= \left(\mathcal{L}_{\overline{e}_\phi} e^{im\phi}\right)\overline{e}_\phi = ime^{im\phi}\overline{e}_\phi = im\overline{e}_{(m)\phi}$$

$$\mathcal{L}_{\overline{e}_\phi} \overline{e}_{(m)j} = \mathcal{L}_{\overline{e}_\phi}\left(e^{im\phi}\, \overline{e}_j\right) = \left(\mathcal{L}_{\overline{e}_\phi} e^{im\phi}\right)\overline{e}_j + e^{im\phi}\mathcal{L}_{\overline{e}_\phi}\left(\overline{e}_j\right)$$

$$= \left(\mathcal{L}_{\overline{e}_\phi} e^{im\phi}\right)\overline{e}_j = ime^{im\phi}\overline{e}_j = im\overline{e}_{(m)j}$$

as desired. Any vector field with axial eigenvalue *m*, i.e.,

$$\mathcal{L}_{\overline{e}_\phi} \overline{V} = im\overline{V}$$

can be expressed as a linear combination of $\{\overline{e}_{(m)\phi}, \overline{e}_{(m)j}\}$ with coefficients that are functions independent of ϕ. In passing, we mention that the Lie draggings along $\overline{e}_\phi$ form a Lie group called $SO(2)$.

3.14. Abstract Lie Groups

3.14.1. Lie Groups

An n-D Lie group is an n-D C^∞ manifold G such that $\forall\, g \in G$, the mappings

$$l_g : G \to G \quad \text{by} \quad h \mapsto l_g(h) = gh \qquad \text{[\textbf{left translation} by } g\text{]}$$

$$r_g : G \to G \quad \text{by} \quad h \mapsto r_g(h) = hg \qquad \text{[\textbf{right translation} by } g\text{]}$$

are diffeomorphisms (C^∞ 1-1 onto maps). These induce corresponding mappings on the tangent spaces:

$$L_g : T_h \to T_{gh} \qquad \text{and} \qquad R_g : T_h \to T_{hg}$$

Of particular interest is the mappings of h in the neighborhood the identity element e.

3.14.2. Lie Algebra

A vector field $\bar{V}$ on G is **left-invariant** if L_g maps $\bar{V}$ at any h to $\bar{V}$ at gh for all g, i.e.,

$$L_g : \bar{V}(h) \mapsto \bar{V}(gh) \qquad \forall\, h, g \qquad \text{(a)}$$

Note that with the help of the group composition, eq(a) is guaranteed if

$$L_g : \bar{V}(e) \mapsto \bar{V}(g) \qquad \forall\, g \qquad \text{(b)}$$

Either L_g or R_g provides a way to compare vectors at different points. For example, left- or right- invariancy is a natural criterion for a **constant** vector field $\bar{V}$ in M. Now, each $\bar{V}(e) \in T_e$ defines a unique left- or right- invariant vector field.

The set of all left- or right- invariant vector fields thus forms a vector space of the same dimension as T_e. [As in §3.10, the coefficients in the linear combinations of these fields must be constants, not functions, on G.] It is easily proved that the Lie bracket of 2 left(right)- invariant fields is also a left(right)- invariant field. Hence, the set of all left- or right- invariant vector fields forms a Lie algebra. By convention, the algebra of the left- invariant fields is called the **Lie algebra** of G and denoted by $\mathcal{L}(G)$. Let $\{\bar{V}_{(i)},\, i = 1, \cdots, n\}$ be a (left-invariant) basis vector fields for $\mathcal{L}(G)$. Closure under the Lie bracket means

$$\left[\bar{V}_{(i)}, \bar{V}_{(j)}\right] = c_{ij}^{k}\, \bar{V}_{(k)} \qquad (3.50)$$

The constants c_{ij}^{k} are called the **structure constants**. It can be shown that they are the components a type $\begin{pmatrix}1\\2\end{pmatrix}$ tensor **C** called the **structure tensor**. Thus, every Lie algebra has a unique structure tensor **C**. The converse of this however not true [see §3.16]. If $c_{ij}^{k} = 0$ for all i, j, and k, then $\mathcal{L}(G)$ is called **abelian**. We shall see that this implies the group G is also abelian.

3.14.3. One-Parameter Subgroups

Consider the integral curve $C(t)$ of a left-invariant vector field $\overline{V}$ that passes through *e* at $t = 0$. Thus, $\overline{V}_e \equiv \overline{V}\big|_e = \frac{d}{dt}\Big|_{t=0}$ and other points on *C* can be obtained by the exponentiation $\exp\left(t\overline{V}\right)$ [see §2.13]. Alternatively, this can be viewed as the diffeomorphism of *G* itself generated by $\overline{V}$ [see §3.1]. Note that $\overline{V}$ is left-invariant so that it is determined entirely by $\overline{V}_e$. To emphasize this point, we write

$$C(t) = g_{\overline{V}_e}(t) = \exp\left(t\overline{V}\right)\Big|_e \tag{3.51}$$

Now, it is easily proved from the definition of exponentiation that

$$\exp\left(t_2\overline{V}\right)\exp\left(t_1\overline{V}\right)\Big|_e = \exp\left[\left(t_2 + t_1\right)\overline{V}\right]\Big|_e$$

Hence,

$$g_{\overline{V}_e}\left(t_1 + t_2\right) = \exp\left[\left(t_1 + t_2\right)\overline{V}\right]\Big|_e \quad = \exp\left(t_2\overline{V}\right)\exp\left(t_1\overline{V}\right)\Big|_e$$

$$= g_{\overline{V}_e}\left(t_2\right) g_{\overline{V}_e}\left(t_1\right) \tag{3.52}$$

so that points on *C* form an abelian group called a **one-parameter subgroup** of *G*. Note that each vector in T_e generates a unique 1-parameter subgroup denoting a C^∞ curve in *G* that passes through *e*. Thus, there is a 1-1 correspondence between these 1-parameter subgroups and the elements of the Lie algebra.

3.15. Examples of Lie Groups

3.15.1. R^n

$\mathbb{R}^n$ is a manifold and an abelian group under vector addition. Hence, it is a Lie group. The 1-parameter subgroups are the 'rays' (straight lines starting from the origin), which are also the left- invariant fields. The Lie algebra is abelian.

3.15.2. $GL(n,K)$

3.15.2.a. Group Manifold

Ther **general linear group** $GL(n,\mathbb{K})$ in *n* dimensions is the group of all invertible $n\times n$ matrices with elements in $\mathbb{K}=\mathbb{R}\ or\ \mathbb{C}$. Group composition is matrix multiplication with the unit matrix *I* as identity *e*. Using the matrix elements as coordinates, we see that the set of all $n\times n$ matrices is $\mathbb{K}^{n^2}$ with $GL(n,\mathbb{K})$ as a submanifold that excludes all points corresponding to matrices of vanishing determinants excluded.

Now, the tangent spaces of $\mathbb{K}^m$ are again $\mathbb{K}^m$. Therefore, the tangent spaces T_P of $GL(n,\mathbb{K})$ are also submanifolds of $\mathbb{K}^{n^2}$ so that all tangent vectors are also $n\times n$ matrices. In fact, $T_P=\mathbb{K}^{n^2}$, i.e., it includes matrices with vanishing determinants. For example, consider the curve through *e* given by

$$P(\lambda)=diag\left(1+e^{\lambda},1,\cdots,1\right)$$

Thus, $\det\left|P(\lambda)\right|=1+e^{\lambda}\neq 0$ for all λ so that the curve is in $GL(n,\mathbb{R})$. Its tangent at *e* is

$$\left.\frac{dP}{d\lambda}\right|_{\lambda=0}=diag(1,0,\cdots,0) \qquad \Rightarrow \qquad \det\left(\left.\frac{dP}{d\lambda}\right|_{\lambda=0}\right)=0$$

Since $T_P=\mathbb{K}^{n^2}$, any matrix can generate a 1-parameter subgroup and hence belongs to the Lie algebra $\mathcal{L}\left[GL(n,\mathbb{K})\right]$.

3.15.2.b. 1-Parameter Subgroups

The 1-parameter subgroup generated by a matrix A is the integral curve $g_A(t)$ of the left-invariant field $\bar{V}$ that is equal to A at e. Setting $t = 0$ at e, we have $\left.\frac{dg_A}{dt}\right|_{t=0} = A$. Using (3.52), we can write

$$g_A(t+\Delta t) = g_A(t)\, g_A(\Delta t)$$

$$\Rightarrow \qquad \frac{dg_A(t)}{dt} = \lim_{\Delta t \to 0} \frac{g_A(t+\Delta t) - g_A(t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{g_A(t)\left[g_A(\Delta t) - I \right]}{\Delta t}$$

Using $g_A(0) = I$, we have

$$g_A(\Delta t) = I + \left.\frac{dg_A}{dt}\right|_0 \Delta t + \cdots$$

so that

$$\frac{dg_A(t)}{dt} = g_A(t) \left.\frac{dg_A}{dt}\right|_0 = g_A(t)\, A \tag{3.53}$$

$$\Rightarrow \qquad g_A(t) = e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n \tag{3.54,5}$$

i.e., the 1-parameter subgroups of $GL(n, \mathbb{K})$ are the exponentiations of arbitrary $n \times n$ matrices, which are called **infintesimal generators** of the subgroup.

3.15.2.c. Component of the Identity

Not every element of $GL(n,\mathbb{R})$ is a member of a 1-parameter subgroup.

This is because a 1-parameter subgroup $g_A(t)$ is a continuous curve that passes through $e = I$ with $\det|I| = +1$. Since $\det|g_A|$ is a continuous function of *t* while $\det|g| \neq 0$ for all $g \in GL(n,\mathbb{R})$, we have $\det|g_A| > 0$. Therefore, every invertible real matrix with negative determinant is not on any 1-parameter subgroup. This means $GL(n,\mathbb{R})$ is a **disconnected** group. In general, elements that belong to a 1-parameter subgroups are called the **component of the identity** of the group.

3.15.2.d. Lie Algebra

Given a tangent vector $\bar{A}_e$ at *e* anf its 1-parameter subgroup $g_{\bar{A}_e}(t)$, the left-translation $f g_{\bar{A}_e}(t)$ of this curve by any matrix $f \in GL(n,\mathbb{K})$ produces a curve of the congruence of the left- invariant vector field $\bar{A}$ generated by $\bar{A}_e$. If *f* is on the curve $g_{\bar{B}_e}(t)$, eq(2.12) gives

$$\left[\bar{A},\bar{B}\right]\Big|_e = \lim_{t\to 0}\frac{1}{t^2}\left\{g_{\bar{A}_e}(t)\,g_{\bar{B}_e}(t) - g_{\bar{B}_e}(t)\,g_{\bar{A}_e}(t)\right\}$$

$$= \lim_{t\to 0}\frac{1}{t^2}\left\{\exp\left(t\bar{A}_e\right)\exp\left(t\bar{B}_e\right) - \exp\left(t\bar{B}_e\right)\exp\left(t\bar{A}_e\right)\right\} \qquad \text{[(3.55) used]}$$

$$= \lim_{t\to 0}\frac{1}{t^2}\left\{\left[I + t\bar{A}_e + \cdots\right]\left[I + t\bar{B}_e + \cdots\right] - \left[I + t\bar{B}_e + \cdots\right]\left[I + t\bar{A}_e + \cdots\right]\right\}$$

$$= \bar{A}_e\bar{B}_e - \bar{B}_e\bar{A}_e \qquad (3.60)$$

which is simply the matrix commutator.

3.15.3. $O(n)$

3.15.4. SU(n)

3.16. Lie Algebras and Their Groups

3.16.1. General Definition of Lie Brackets

Every Lie group G has its Lie algebra $\mathfrak{G}$. Now, every element $g \in G$ is the image of e under the left- translation generated by g. Also, every vector $\overline{V}_e \in T_e$ generates a unique vector field $\overline{V}$ in $\mathfrak{G}$. Therefore, every g is on one curve of each of the left- invariant congruences. The question is whether it is possible to reconstruct G given $\mathfrak{G}$.

To answer the question, we first generalize the definition of a Lie bracket and hence that of the Lie algebra. Thus, the Lie bracket is defined as an internal binary operator on a vector space V on $\mathbb{K}$,

$$V \times V \to V \qquad \text{by} \qquad \left(\overline{A},\overline{B}\right) \mapsto \left[\overline{A},\overline{B}\right]$$

such that for all $\overline{A},\overline{B} \in V$ and $a,b \in \mathbb{K}$,

1. $\left[a\overline{A},\overline{B}\right] = a\left[\overline{A},\overline{B}\right]$ and $\left[\overline{A},b\overline{B}\right] = b\left[\overline{A},\overline{B}\right]$. (bilinearity)

2. $\left[\overline{A},\overline{B}\right] = -\left[\overline{B},\overline{A}\right]$. (antisymmetry) (3.68)

3. $\left[\overline{A},\left[\overline{B},\overline{C}\right]\right] + \left[\overline{B},\left[\overline{C},\overline{A}\right]\right] + \left[\overline{C},\left[\overline{A},\overline{B}\right]\right] = 0$. (Jacobi identity) (3.69)

For example, the cross product of vectors in $\mathbb{K}^3$ defines a Lie bracket:

$$\left[\overline{a},\overline{b}\right] \equiv \overline{a} \times \overline{b} \qquad \forall\ \overline{a},\overline{b} \in \mathbb{K}^3 \tag{3.70}$$

3.16.2. Coverings

We now state but not prove the following theorems:

1. Every Lie algebra is the Lie algebra of one and only one *simply- connected* Lie group.
2. Any Lie group that is multiply- connected can be *covered* by the simply- connected one belonging to the same Lie algebra. The covering is then a homomorphism between these groups.

Note that

1. A manifold is simply- connected if every closed curve in it can be smoothly shrunk into a point.
2. A connected manifold M **covers** another manifold N if there is a map

   $$\pi : M \to N$$

   that is onto and such that for any neighborhood U of a point $P \in N$,

   $$\pi^{-1}(U) = \bigcup_i \pi_i^{-1}(U)$$

   where $\pi^{-1}(U)$ is the inverse image of U and π_i^{-1} denotes the inverse function for the *i*th 1-1 branch of π.

For example, the real line $\mathbb{R}$ covers the unit circle S^1 an infinite number of times by the map

$$\pi : \mathbb{R} \to S^1 \qquad \text{by} \qquad x \mapsto \pi(x) = (\cos x, \sin x)$$

Thus, every interval $[x, x+2\pi]$ in $\mathbb{R}$ is mapped onto S^1

3.16.3. SU(2)

Consider the set H of matrices of the form

$$\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \qquad (3.72)$$

We leave as exercise the proofs of the following:

1. The set $H - \left\{ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}$ is a group under matrix multiplication. It is called the

   $GL\left(2,\mathbb{C}\right)$.

2. H is a 4-D real vector space under matrix addition. One basis is

   $$J_1 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \qquad J_2 = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad J_3 = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \qquad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

3. Writing $A \in H$ as

   $$A = 2\alpha_1 J_1 + 2\alpha_2 J_2 + 2\alpha_3 J_3 + \alpha_4 I$$

   where $\left\{\alpha_j\right\}$ are real. We have $A \in SU(2)$ iff

   $$\sum_{j=1}^{4} \alpha_j = 1 \qquad (3.73)$$

4. There is a 1-1 onto mapping from $SU(2)$ to the spherical surface S^3.

Since S^3 is simply- connected, so is $SU(2)$.

3.16.4. Exponentiations for $SU(2)$

Elements of $SU(2)$ can be obtained by the exponentiation of the elements of $\mathcal{L}[SU(2)]$.

For example, using

$$J_1 = \frac{i}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad J_1^2 = \left(\frac{i}{2}\right)^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \left(\frac{i}{2}\right)^2 I$$

$$J_1^3 = \left(\frac{i}{2}\right)^3 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \left(\frac{i}{2}\right)^2 J_1$$

so that

$$J_1^{2n+1} = \left(J_1^2\right)^n J_1 = \left(\frac{i}{2}\right)^{2n} J_1$$

$$J_1^{2n} = \left(J_1^2\right)^n = \left(\frac{i}{2}\right)^{2n} I \qquad n = 0,1,\cdots$$

the exponentiation of J_1 gives

$$\exp\left(tJ_1\right) = \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!} J_1^{2n} + \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!} J_1^{2n+1}$$

$$= \sum_{n=0}^{\infty} \frac{1}{(2n)!}\left(\frac{it}{2}\right)^{2n} I + 2\sum_{n=0}^{\infty} \frac{i^{2n}}{(2n+1)!}\left(\frac{t}{2}\right)^{2n+1} J_1$$

$$= I \cos\frac{t}{2} + 2J_1 \sin\frac{t}{2}$$

$$= \begin{pmatrix} \cos\frac{t}{2} & i\sin\frac{t}{2} \\ i\sin\frac{t}{2} & \cos\frac{t}{2} \end{pmatrix} \qquad (3.74)$$

3.16.5. Exponentiations for $SO(3)$

For the L_1 generator of $SO(3)$, we have

$$L_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad\qquad L_1^2 = -\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = -I_1$$

$\Rightarrow$

$$L_1^{2n+1} = \left(L_1^2\right)^n L_1 = (-)^n L_1 \quad = (-)^n \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$L_1^{2n} = \left(L_1^2\right)^n = (-)^n I_1$$

so that

$$\begin{aligned} \exp(sL_1) &= I + \sum_{n=0}^{\infty} \frac{s^{2n+1}}{(2n+1)!}(-)^n L_1 + \sum_{n=1}^{\infty} \frac{s^{2n}}{(2n)!}(-)^n I_1 \\ &= I + L_1 \sin s + I_1 \cos s \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos s & -\sin s \\ 0 & \sin s & \cos s \end{pmatrix} \end{aligned} \tag{3.75}$$

3.16.6. *SU*(2) Covers *SO*(3)

The exponentiations given by (3.74-5) suggests a mapping

$$\pi : SU(2) \to SO(3)$$

by

$$\exp(tJ_1) \mapsto \exp(tL_1)$$

i.e., $$\begin{pmatrix} \cos\frac{t}{2} & i\sin\frac{t}{2} \\ i\sin\frac{t}{2} & \cos\frac{t}{2} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix} \qquad (3.76)$$

Now, the distinct elements in $\exp(tL_1)$ are represented by any interval

$t \in [a, a+2\pi]$ of length 2π and those in $\exp(tJ_1)$ by $t \in [a, a+4\pi]$ of length 4π.

Thus, $\exp(tJ_1)$ **doubly covers** $\exp(tL_1)$. Similar conclusions can be drawn for

the other generators so that the map

$$\pi : SU(2) \to SO(3)$$

by

$$\exp\left(\sum_{j=1}^{3} t_j J_j\right) \mapsto \exp\left(\sum_{j=1}^{3} t_j L_j\right) \qquad (3.77)$$

is a double covering of *SO*(3) by *SU*(2).

3.16.7. Topologies

As mentioned in §3.16.3, the manifold $SU(2)$ is diffeomorphic to the simply connected S^3 and hence share the same global topology. A one-parameter subgroup $\exp(tJ_1)$ of $SU(2)$ begins at e with $t=0$ and returns to it at $t=4\pi$, thus tracing out a great circle on S^3. The points labelled t and $t+2\pi$ are diametrically opposite to each other. Since they are mapped to the same point in $SO(3)$, we see that $SO(3)$ is a hemisphere of S^3. The manifold $SO(3)$ is not simply connected because the great circle forming its boundary cannot be shrunk since points on it must be pairwise diametrically opposite. Note that locally, $SU(2)$ and $SO(3)$ are identical so that they have the same Lie algebra.

The Lie algebra using the cross product as Lie bracket [see (3.70)] can be associated with either group $SU(2)$ or $SO(3)$. In classical mechanics, cross products in $\mathbb{R}^3$ are usually associated with $SO(3)$ so that elements of the subgroup $\exp(\theta L_j)$ correspond to rotations about the x^j axis. However, this identification of rotations with vectors works only in $\mathbb{R}^3$. For example, vectors in $\mathbb{R}^4$ are 4-D but the rotation group $SO(4)$ is 6-D so that no such identification is possible. In quantum mechanics, associating $SU(2)$ with $\mathbb{R}^3$ allows the assignment of the spin to a vector in $\mathbb{R}^3$ even though the spin is not an element of the tangent space of $\mathbb{R}^3$.

Finally, we mention that the algebra of an abelian group must be abelian.

3.17. Realizations and Representations

3.17.1. Definitions

Mathematically, a group, say *SO*(3), is an **abstract group** defined entirely by its group operation and manifold topology. However, in applications in physics, the importance of a group is its *action* on physical quantities. For example, elements of *SO*(3) are associated with rotations of objects in 3-D space. Such an association is called a realization. Let $T(M)$ be the set of operators on some space M. A **realization** of a group *G* is a map

$$T : G \to T(M) \qquad \text{by} \qquad g \mapsto T(g)$$

such that the group properties are preserved, i.e.,

1. $T(e) = I$ where *I* is the identity transformation on *M*.

2. $T(g^{-1}) = \left[T(g)\right]^{-1}$.

3. $T(g) \circ T(h) = T(gh)$ where $\circ$ is the composition in $T(M)$.

A realization is **faithful** if *T* is 1-1, i.e.,

$$g \neq h \qquad \Rightarrow \qquad T(g) \neq T(h)$$

If *M* is a vector space and *T*(*g*) are linear transformations, the realization is called a **representation**.

3.17.2. Example 1

Consider the unit sphere S^2 in $\mathbb{R}^3$ given by $x^2 + y^2 + z^2 = 1$. A rotation by θ about the *x*-axis maps a point (x, y, z) to another (x', y', z') according to

$$\begin{aligned} x' &= x \\ y' &= y\cos\theta - z\sin\theta \qquad (3.78) \\ z' &= y\sin\theta + z\cos\theta \end{aligned}$$

so that

$$x'^2 + y'^2 + z'^2 = 1$$

This transformation can be associated with an element in the subgroup $\exp(\theta L_1)$ of $SO(3)$. If we treat it as a transformation of S^2 into itself, we have a realization of $SO(3)$ since S^2 is not a vector space. On the other hand, if we treat it as transformations in $\mathbb{R}^3$, which is a vector space, we have a representation of $SO(3)$.

The fact that, originally, we have used rotations in $\mathbb{R}^3$ to define $SO(3)$ illustrates a useful technique. Thus, a group is first defined by a faithful realization or representation so that its properties can be studied in concrete terms. Afterwards, the group is regarded as abstract to allow for other useful realizations and representations.

3.17.3. Example 2

Every group *G* has at least 2 faithful realizations: the left and right translations of itself. The left translations give rise to the **progressive** (**principal**) **realization** and the right translations to the **retrograde realization**.

3.17.4. Example 3

Each Lie group G has a representation using its elements as linear transformations on its own Lie algebra. This is called the **adjoint representation** and is defined as follows. First, we define the **adjoint realization** of the group G by the mappings

$$I_g : G \to G \qquad \text{by} \qquad h \mapsto I_g(h) = ghg^{-1} \qquad \forall\ g, h \in G$$

which is not necessarily faithful. For example, if G is abelian, I_g is the identity map $h \mapsto h$ for all g. The mappings I_g are **inner automorphisms** of G. Note that $I_g(e) = geg^{-1} = e$ for all g so that every curve through e is mapped to another curve through e. Thus, I_g induces another map

$$Ad_g : T_e \to T_e$$

called the **adjoint transformation** of T_e induced by g. Given a 1-parameter subgroup $\exp(t\bar{X})$ where $\bar{X} \in T_e$, its image under I_g is another 1-parameter subgroup with composition

$$\left(gfg^{-1}\right)\left(ghg^{-1}\right) = g\,fhg^{-1}$$

This defines the action of Ad_g as

$$I_g\left[\exp\left(t\bar{X}\right)\right] = \exp\left[tAd_g\left(\bar{X}\right)\right] \qquad (3.79)$$

If g is also a member of the 1-parameter subgroup $g(s) = \exp\left(s\bar{Y}\right)$, it is left as an exercise to show that

$$Ad_{g(s)}\left(\bar{X}\right) = \exp\left(s\mathcal{L}_{\bar{Y}}\right)\bar{X} \qquad (3.80)$$

3.18. Spherical Symmetry, Spherical Harmonics and Representations of the Rotation Group

3.18.1. Preliminary

The following is a collection of definitions that are used later.

1. A **sequence** $\{x_n\}$ of points on a manifold *M* is a mapping

   $$\mathbb{N} \to M \qquad \text{by} \qquad n \mapsto x_n$$

2. A sequence $\{x_n\}$ **converges** to $x \in M$ if

   $$\forall\, N(x),\ \exists\, n_0 \in \mathbb{N} \text{ such that } x_n \in N(x) \quad \forall\, n \geq n_0$$

   where *N*(*x*) is a neighborhood of *x*.

3. A **Cauchy sequence** is a sequence such that

   $$\forall\, N(0),\ \exists\, n_0 \in \mathbb{N} \text{ such that } x_m - x_n \in N(0) \quad \forall\, n,m \geq n_0$$

4. A **topological vector space** is a manifold which is also a vector space.
5. A topological vector space is **complete** if every Cauchy sequence in it converges to some point in it.
6. Let *V* be a linear space on $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. A mapping

   $$V \times V \to \mathbb{C} \qquad \text{by} \qquad (x, y) \to \langle x | y \rangle$$

   is a **sesquilinear mapping** if for all $x, y, z \in V$ and $a, b \in \mathbb{C}$,

   $$\langle x | y \rangle = \langle y | x \rangle^*$$

   $$\langle ax + by | z \rangle = a^* \langle x | z \rangle + b^* \langle y | z \rangle$$

7. A sesquilinear mapping is **positive** if $\langle x | x \rangle \geq 0 \quad \forall\, x$.
8. A sesquilinear mapping is **strictly positive** if it is positive with

   $$\langle x | x \rangle = 0 \text{ iff } x = 0.$$

9. A **seminorm** is a mapping

   $$V \to \mathbb{R} \qquad \text{by} \qquad x \mapsto \|x\|$$

   such that for all $x, y \in V$ and $a, b \in \mathbb{K}$, we have,

   $$\|x + y\| \leq \|x\| + \|y\| \qquad \text{(triangular inequality)}$$

   $$\|ax\| = |a|\,\|x\|$$

If furthermore,

$$\|x\| = 0 \qquad \text{iff} \qquad x = 0$$

it is a **norm**.

10. A complete normed vector space is called a **Banach space**.
11. A linear space with a strictly positive sesquilinear mapping is called a **pre–Hilbert space** & the mapping the **inner product**.
12. The inner product induces a **norm** $\|x\| = \sqrt{\langle x | x \rangle}$.
13. A complete pre-Hilbert space is called a **Hilbert space**.
14. The space of all functions *f* on *M* such that $|f|^p$ is integrable is called $L^p(M)$.

3.18.2. Spherical Symmetry

A manifold M is **spherically symmetric** if the Lie algebra of its Killing vector fields has a subalgebra $so(3) \equiv \mathcal{L}[SO(3)]$. Consider the Hilbert space $H(S^2)$ on S^2 which is the linear (function) space $L^2(S^2)$ of all square- integrable complex functions on S^2. A vector in $L^2(S^2)$ is a function f with norm

$$\|f\| = \sqrt{\int_{S^2} dx\; f^*(x)\, f(x)}$$

where $f(x)$ is the component of the vector f in the basis consists of eigenfunctions of the position operator. Note that $L^2(S^2)$ is a vector space of infinite dimensions.

The realization of $g \in SO(3)$ as a mapping

$$R(g): S^2 \to S^2 \qquad \text{by} \qquad x \mapsto x' = R(g)\, x$$

induces a mapping

$$\mathcal{R}(g): L^2(S^2) \to L^2(S^2) \qquad \text{by} \qquad f \mapsto f' = \mathcal{R}(g)\, f$$

With suitable definitions, $\mathcal{R}(g)$ can be made to be a representation of $SO(3)$ on $L^2(S^2)$. Since $L^2(S^2)$ is infinite dimensional, so is $\mathcal{R}(g)$.

There are also finite dimensional subspaces of $L^2(S^2)$ that are invariant under $\mathcal{R}(g)$ $\forall\, g$. They form representations of finite dimensions. Invariant subspaces that contain no smaller invariant subspaces give **irreducible representations** (IRs).

Let $\{ f_i \,; i = 1, \cdots, N \}$ be a basis for such an invariant subspace $V \subset L^2(S^2)$, then

$$f = a^i f_i \qquad \forall\; f \in V$$

$$\mathcal{R}(g)\, f = \mathcal{R}(g)\left[a^i f_i \right] = a^i \mathcal{R}(g)\, f_i = f' = b^i f_i$$

which implies

$$\mathcal{R}(g)\, f_i = f_j\, g^j_i$$

such that

$$a^i f_j g_i^j = b^i f_i = b^j f_j \qquad \Rightarrow \qquad b^j = g_i^j a^i$$

The matrix $m_{ij} = g_j^i$ is called the representation of g in V.

The basis functions of finite dimensional IRs for $SO(3)$ are the spherical harmonics Y_{lm}. (see Tung for the actual construction).
Thus, each invariant subspaces V_l of $L^2(S^2)$ is characterized by an integer $l \geq 0$ and has dimension $2l+1$. The set $Y_l = \{Y_{lm} \, ; m = -l, \cdots, l\}$ is a basis of V_l. The spherical harmonics are complete in the sense that

$$Y = \bigcup_{l=0}^{\infty} Y_l \quad = \{Y_{lm} \, ; l = 0, 1, \cdots ; m = -l, \cdots, l\}$$

is a basis for $L^2(S^2)$.

Since $R(g) = \exp\left(t^i \overline{l}_i\right)$ and $\mathcal{R}(g) = \exp\left(t^i \overline{\ell}_i\right)$, V_l is invariant under $\mathcal{R}(g)$ iff it is invariant under $\left\{\overline{\ell}_i \, ; i = x, y, z\right\}$. This is guaranteed if the basis Y_l is invariant under the generators $\left\{\overline{\ell}_i \, ; i = x, y, z\right\}$, i.e.,

$$\overline{\ell}_i Y_{lm} = \mathcal{L}_{\overline{\ell}_i} Y_{lm} \quad = \sum_k c_k Y_{lk} \qquad \forall \, i = x, y, z \quad \text{and} \quad m = -l, \cdots, l$$

For example, if $l = 0$,

$$Y_0 = \{Y_{00} = 1\}$$

$$\overline{\ell}_i \left(Y_{00}\right) = \mathcal{L}_{\overline{\ell}_i} Y_{00} = \overline{\ell}_i \left(1\right) = 0 = 0 \cdot Y_{00} \qquad \forall \, i = x, y, z$$

For $l = 1$,

$$Y_1 = \{Y_{11}, Y_{10}, Y_{1-1}\} \quad = \left\{\sqrt{\frac{3}{8\pi}} \sin\theta \, e^{i\phi}, \sqrt{\frac{3}{4\pi}} \cos\theta, \sqrt{\frac{3}{8\pi}} \sin\theta \, e^{-i\phi}\right\}$$

$$= \left\{\sqrt{\frac{3}{8\pi}} \left(x + iy\right), \sqrt{\frac{3}{4\pi}} z, \sqrt{\frac{3}{8\pi}} \left(x - iy\right)\right\}$$

Thus

$$\overline{l}_x \left(Y_{l\pm1}\right) = \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y}\right) \left[\sqrt{\frac{3}{8\pi}} \left(x \pm iy\right)\right]$$

$$= \sqrt{\frac{3}{8\pi}} \left(\mp zi \right) \quad = \mp \frac{i}{\sqrt{2}} Y_{10}$$

$$\overline{l}_z \left(Y_{l\pm1} \right) = \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \left[\sqrt{\frac{3}{8\pi}} \left(x \pm iy \right) \right]$$

$$= \sqrt{\frac{3}{8\pi}} \left(\pm ix - y \right) \quad = \pm i \sqrt{\frac{3}{8\pi}} \left(x \pm iy \right) \quad = \pm i \, Y_{1\pm1}$$

$l = 1$ is the smallest faithful representation of $SO(3)$. It is usually called the **fundamental representation**. Obviously, it is equivalent to the defining 3×3 special orthogonal matrix representation. (see Ex 3.26, Schultz)

3.18.3. $SO(3)$ and $SU(2)$

Since $\pi : SU(2) \to SO(3)$ is a 2-1 mapping, i.e., there are 2 elements $u, u' \in SU(2)$

that maps into the same $g \in SO(3)$, therefore, any representation $R(g)$ of $g \in SO(3)$ defines a (unfaithful) representation S for $SU(2)$ such that

$$S(u) = S(u') = R[\pi(u)] = R[\pi(u')]$$

Other representations, say T , of $SU(2)$ such that

$$T(u) \neq T(u') \qquad \text{even if} \qquad \pi(u) = \pi(u')$$

are called **double valued representations** of $SO(3)$.

IRs of $SU(2)$ are characterized by an index k which is either a whole or a half integer. If k is an integer, they are representations of $SO(3)$ with the same index ($l = k$). If k is a half- integer, they're double- valued representations of $SO(3)$.

The 2×2 special unitary matrices used to define $SU(2)$ form a representation with $k = \frac{1}{2}$. It is the smallest faithful representation & often called the spin $\frac{1}{2}$ representation. Elements of the vector space are called **spinors**.

3.19. Bibliography