Maithripala D. H. S.,
Dept. of Mechanical Engineering,
Faculty of Engineering,
University of Peradeniya,
Sri Lanka.

Lecture Notes on Advanced Mechanics

My Attempt at Learning

April 2, 2020

Preface

This is a compilation of notes on Physics as I go about learning it the hard and only way I know.

Peradeniya, Sri Lanka, *D. H. S. Maithripala*
April 2, 2020

Contents

Chapter 1
Einstein's Theory of Special Relativity

Recall that Galilean mechanics was founded upon the assumption that time intervals, distance between points, and the mass of a particle, were observer independent. In special relativity we will be a little less restrictive and assume:

Axiom 1.1 Fundamental Assumptions of Relativistic Space-Time: *There exists special observers, referred to as inertial observers, all of whom agree that*

(i) space-time is 4-dimensional, continuous, homogeneous, and isotropic,
(ii) light travels in a straight line,
(iii) the speed of light is an inertial observer independent quantity.

In special relativity, similar to the case of Galilean mechanics, different observers are represented by different coordinate (measurement) systems for space-time. The assumption that light moves in a straight line and that the observed (measured) speed of light is constant allows an inertial observer to synchronise clocks at different points in space to read the 'same' time as the one held by the observer. This same assumption allows an inertial observer to define lengths by measuring the distances that light travels. The notion of straight lines and distances gives rise to the notion of orthogonality. Thus these two assumptions imply that space as observed by the inertial observer is Euclidean[1]. This notion allows the inertial observer to setup three mutually perpendicular axis in space and identify a point in space with three numbers that correspond to the perpendicular distances from the point to each of the three axis of the frame. Therefore the synchronised clocks and the reference orthonormal spatial frame allows the inertial observer to identify a space time event A occurring at the spacial point P by the four numbers $(t,\mathbf{x}) \in \mathbb{R}^4$ where $t \in \mathbb{R}$ is the time measured by the clock that the observer has and $\mathbf{x} \in \mathbb{R}^3$ is the representation of the point P in the orthonormal spatial frame. Consider two events occurring at the spatial points P and Q. Denote by $(t_P,\mathbf{x_P})$ the representation of the event at P and denote by $(t_Q,\mathbf{x_Q})$ the representation of the event at Q as observed by the inertial observer. Since the first two assumptions imply that the space observed by the inertial observer is Euclidean, we can define the *spatial distance* between the two events at P and Q,

[1] A space where all the notions of Euclidean geometry hold is called an Euclidean space.

by virtue of the Pythagorean theorem, to be equal to the Euclidean distance between the two points $\mathbf{x_P} \in \mathbb{R}^3$ and $\mathbf{x_Q} \in \mathbb{R}^3$. That is, the spatial distance observed by the inertial observer between the events occurring at P and Q to be given by the number $||\mathbf{x_Q} - \mathbf{x_P}||$. Note that we will use boldfaced symbols to denote special elements.

In summary the first two assumptions of the principles of relativity imply that every inertial observer can define a preferred globally defined coordinate system for 4D-space-time that gives the spatial distance between two events occurring at two different spatial points by the Euclidean distance between the two points. These preferred coordinate (measurement) systems are usually called *Lorentz* coordinates and observers who can define such coordinates[2] are called *inertial* observers. In general a different inertial observer will have a different clock and a different orthonormal frame and thus would define a different Lorentz coordinate system that will assign to a space-time event P an element $(t', \mathbf{x}') \in \mathbb{R} \times \mathbb{R}^3 \equiv \mathbb{R}^4$.

Axiom 1.2 Relativity Principle: *A fundamental hypothesis of relativistic mechanics is that laws of nature should be independent of the inertial observer or in other words the mathematical fundamental expressions that describe the laws of physics should be the same in any Lorentz coordinate system.*

In order to study concepts, and properties invariant of inertial observers we need to know how each of these measurement systems (Lorentz coordinates) are related to each other. For simplicity, let us first consider the case of one dimensional space and time (that is let us consider 2D space-time). Consider two inertial observers **e** and **b**. Let the Lorentz coordinates of a particular event in space-time be given by $(t, \mathbf{x})$ in the Lorentz coordinates of **e** and by $(t', \mathbf{x}')$ in the Lorentz coordinates of **b**. The assumption that every observer sees that space-time is continuous and homogeneous implies that the two coordinates are necessarily related by a linear transformation of the form

$$\begin{aligned} t' &= \gamma t + \beta \mathbf{x}, \\ \mathbf{x}' &= \alpha t + \eta \mathbf{x}, \end{aligned}$$

where $\alpha, \beta, \gamma, \eta$ are constants where without loss of generality the origin of both coordinate systems are chosen to coincide.

Considering the representation of the origin O' of the spatial frame of **b** in the spatial frame of **e** we have that

$$\alpha = -\eta \mathbf{v}.$$

where $\mathbf{v}$ is the velocity of O' as observed in the **e** coordinates and hence is a constant. At the same time the isotropy of space implies that the velocity of the origin of **e**, O , observed in the **b** frame must be $-\mathbf{v}$. This means that inertial observers move at a constant speed with respect to each other. Thus we must have

$$\frac{\alpha}{\gamma} = -\mathbf{v}.$$

[2] That is, those observers who confirm that the assumptions of Axiom-1.1 are true.

Which gives $\eta \mathbf{v} = \gamma \mathbf{v}$ and $\alpha = -\gamma \mathbf{v}$ and thus the transformations become

$$t' = \gamma t + \beta \mathbf{x},$$
$$\mathbf{x}' = -\gamma \mathbf{v} t + \gamma \mathbf{x}.$$

Recall that in Galilean Physics we assume that time is invariant for all observers and hence we must have $\beta = 0$, and $\gamma = 1$, thus

$$t' = t, \tag{1.1}$$
$$\mathbf{x}' = \eta \mathbf{x} - \mathbf{v} t. \tag{1.2}$$

In Galilean mechanics spacial distance differences are also assumed to be invariant of observers. Thus necessarily $\eta = 1$ (in the general 3D case it must be an orthonormal transformation).

We have seen that coordinate systems for Galilean space time satisfying these relationships were also called inertial. Notice that since $\beta = 0$ the speed of light observed by all Galilean inertial observers are not the same.

In special relativity we have abandoned the assumption that time is invariant for all observers and instead have assume that the speed of light is the same for all observers. Let $(t, \mathbf{x})$ be the representation of the event of a light particle in the Lorentz coordinates of the inertial observer $\mathbf{e}$ and let $(t', \mathbf{x}')$ be the representation of the same event of the light particle in the Lorentz coordinates of the inertial observer $\mathbf{b}$. For convenience, in the following we assume that $\mathbf{b}$ is moving at a constant speed, v, in the $\mathbf{e}_i$ direction with respect to $\mathbf{e}$ and that the light particle is also moving in the $\mathbf{e}_i$ direction. Then the invariance implies that

$$\frac{d\mathbf{x}_i}{dt} = \frac{d\mathbf{x}_i'}{dt'} = c,$$

$\mathbf{x}_j = \mathbf{x}_j'$, and $\mathbf{x}_k = \mathbf{x}_k'$ where i, j, k is any permutation of $1, 2, 3$. Thus from the transformations we have

$$\frac{d\mathbf{x}'}{dt'} = \frac{\gamma(-v + c)}{\gamma + c\beta_i} = c,$$

which gives

$$\beta_i = -\frac{v\gamma}{c^2}.$$

Since $\mathbf{e}$ and $\mathbf{b}$ are parallel and $\eta e_i = \gamma e_i$ we then have that

$$t' = \gamma \left(t - \frac{v}{c^2} \mathbf{x}_i \right), \tag{1.3}$$
$$\mathbf{x}_i' = \gamma \left(-vt + \mathbf{x}_i \right), \tag{1.4}$$
$$\mathbf{x}_j' = \mathbf{x}_j, \tag{1.5}$$
$$\mathbf{x}_k' = \mathbf{x}_k. \tag{1.6}$$

The particular type of transformation given above is usually referred to as a *boost in the* i^{th} *spatial direction*.

Considering the inverse transformation of the boost in the i^{th} direction and the assumption that 4D-space-time is isotropic we have

$$\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}. \tag{1.7}$$

With a little more effort but following similar arguments as before it can be shown that a general Lorentz transformation takes the form,

$$\begin{bmatrix} t' \\ \mathbf{x}' \end{bmatrix} = \begin{bmatrix} \gamma & -\frac{\gamma}{c^2}\mathbf{v}^T \\ -\gamma\mathbf{v} & ((\gamma-1)\frac{\mathbf{v}\mathbf{v}^T}{||\mathbf{v}||^2} + I) \end{bmatrix} \begin{bmatrix} t \\ \mathbf{x} \end{bmatrix}. \tag{1.8}$$

One can show that the set of all such transformations, $\Lambda(\mathbf{v})$, that relate two Lorentz coordinate systems form a group called the *Lorentz group*.

Let us try to summarise what Axiom-1.1 means. It means that there exists special class of globally defined coordinates (observers) for 4D-space-time called Lorentz coordinates (inertial observers) such that the motion of a light particle represented using any one of these coordinates correspond to a constant velocity straight line motion. In a physical sense, different Lorentz coordinate systems are shown to correspond to different observers moving at constant velocity with respect to each other without rotation. It also says that observers accelerating with respect to inertial observers will not see that light particles move in a straight line at constant velocity.

We point out an un-intuitive consequence of Axiom-1.1. Consider again the two clocks used by the two inertial observers $\mathbf{e}$ and $\mathbf{b}$; where one is at O in space and the other is at O' in space. Let $v = ||\mathbf{v}||$. We saw that the point O has the representation $(\gamma t, -\mathbf{v}\gamma t)$ in the $\mathbf{b}$ coordinates. Thus an inertial observer $\mathbf{b}$ sees that the clock at O is going slow. At the same time the point O' has the representation $(\gamma t', \mathbf{v}\gamma t')$ in the Lorentz co-ordinates of $\mathbf{e}$. Thus the inertial observer $\mathbf{e}$ also sees that the clock at O' is going slow. That is moving clocks go slower! This phenomena is referred to as *time dilation* and is simply a consequence of the assumptions of relativity stated in Axiom-1.1.

We have seen that the laws of Galilean mechanics were a statement about invariance of the laws of nature governing the motion of objects that are moving at speeds sufficiently smaller than the speed of light. It was stated in the form of a conservation law for the total linear momentum of a collection of interacting but isolated set of particles. The mass of each particle and the fundamental interactions between particles were assumed to be observer independent in Galilean mechanics. Similarly in relativistic mechanics we will look for principles and relationships that are the same in all relativistic inertial frames. Such principles will be called *Laws of Nature*.

In this search the first physical principle that follows immediately from the properties of Lorentz transformations is:

The Inertia Property: If a particle is observed to move in a straight line at a constant speed in a particular Lorentz coordinate system, it will also be observed to move in a straight line at a constant speed in any other Lorentz coordinate system.

Next it can be shown that the Minkowski Pseudo-Riemannian metric

$$ds^2 = dt \otimes dt - \frac{1}{c^2}(dx^1 \otimes dx^1 + dx^2 \otimes dx^2 + dx^3 \otimes dx^3). \tag{1.9}$$

is invariant under Lorentz transformations. Thus it can be used as the inertial observe independent (invariant) measure of distance between events in space-time[3].

Consider an inertial observer $\mathbf{e}$ and the corresponding Lorentz coordinates $(t, \mathbf{x})$ and the path of a particle $q(t) = (t, \mathbf{x}(t))$ parameterized by the co-ordinate time parameter t. This is called a world line. Then the velocity vector with respect to t is $\dot{q} = (1, \dot{\mathbf{x}})$ where $\mathbf{v} = \dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt}$ is the velocity measured in the $\mathbf{e}$ frame. Different inertial observers will observe different velocity vectors because their co-ordinate times are different. However all inertial observers will observe the same arc-length since (1.9) is invariant under Lorentz transformations. Similar to the case of 3D- Euclidean space, the arc-length parameter τ is defined to be the square-root of the pull-back of ds^2 to the world-line. That is,

$$d\tau \triangleq \left(q^* ds^2\right)^{\frac{1}{2}} = \left(1 - \frac{v^2}{c^2}\right)^{\frac{1}{2}} dt.$$

Thus we have that

$$\frac{dt}{d\tau} = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} = \gamma.$$

Since the arc-length is invariant for all inertial observers it is called the *proper time parameter*. It corresponds to the time kept by a clock moving with the particle.

When the world line is parameterized by the arc-length parameter τ as $q(\tau) = (t(\tau), \mathbf{x}(\tau))$ the velocity vector with respect to the arc-length parameter τ is given by $U(\tau) \triangleq \dfrac{dq}{d\tau} = (\frac{dt}{d\tau}, \frac{dt}{d\tau}\mathbf{v})$ where $\mathbf{v} = \frac{d\mathbf{x}}{dt}$ is the velocity measured by $\mathbf{e}$. Thus we have $U(\tau) = \gamma(1, \mathbf{v})$. We verify that

$$<< U, U >>= \left(\frac{dt}{d\tau}\right)^2 \left(1 - \frac{v^2}{c^2}\right) = 1,$$

as should be the case when any trajectory is parameterized by arc-length.

Denote by m, the mass of a particle as observed by $\mathbf{e}$. Motivated by the Galilean case we define a *linear momentum 4-vector* in each Lorentz coordinate system by the relationship

$$P \triangleq m(1, \mathbf{v}) = (m, m\mathbf{v}).$$

The magnitude of the momentum 4-vector with respect to the Minkowski metric is $||P|| = m/\gamma = m_r$. In special relativity we assume the following:

[3] Space-time equipped with this metric is called Minkowsky space.

Axiom 1.3 *The magnitude of the momentum 4-vector, $||P|| = m_r$, is a constant for all inertial observers.*

The quantity m_r is called the *rest mass* of the particle. The mass observed by $\mathbf{e}$, that is given by $m = m_r\gamma$, is called the *relativistic mass* observed by the inertial observer $\mathbf{e}$. Thus in special relativity we assume that the rest mass, instead of the observed mass, is an inertial observer independent quantity. From the assumption of the existence of a rest mass it follows that

$$P = (m, m\mathbf{v}) = (m_r\gamma, m_r\gamma\mathbf{v}) = m_r U.$$

Consider a set of isolated N number of particles each of rest mass m_{ri}. Consider the description of the motion of these particles in a Lorentz co-ordinate system $\mathbf{e}$. The total momentum of the system of particles is given by

$$\bar{P} = \sum_{i=1}^{N} P_i = \sum_{i=1}^{N} m_{ri} U_i = \sum_{i=1}^{N} (m_i, m_i \mathbf{v}_i) = \left(\sum_{i=1}^{N} m_i, \sum_{i=1}^{N} m_i \mathbf{v}_i\right).$$

The total momentum of the set of particles expressed in a different Lorentz coordinate system moving at a relative velocity $\mathbf{v}$ (thus related by the Lorentz transformation $\Lambda(v)$) will take the form

$$\bar{P}' = \sum_{i=1}^{N} P_i' = \sum_{i=1}^{N} m_{ri} U_i' = \sum_{i=1}^{N} m_{ri} \Lambda(\mathbf{v}) U_i = \Lambda(\mathbf{v}) \bar{P}.$$

Thus if the total momentum of the system of interacting particles is observed to be conserved in one Lorentz frame it should be observed to be constant in any other Lorentz frame. From the first expression we se that conservation of momentum of a collection of interacting particles imply that the total mass of the particles and the total spatial linear momentum of the particles are conserved. Experiments indicate that this is true if the interacting particles are isolated. Thus in special relativity, in agreement with experiment, it is postulated that:

Axiom 1.4 Conservation of momentum: *The total 4-momentum of a system of isolated and interacting particles is conserved in any Lorentz coordinate systems.*

It is interesting to note that this axiom unites conservation of mass and spatial momentum into one single principle.

The principle of conservation of momentum in each Lorentz frame implies that in any Lorentz coordinate system

$$(0, \mathbf{0}) = \frac{d\bar{P}}{dt} = \left(\sum_{i=1}^{N} \frac{dm_i}{dt}, \sum_{i=1}^{N} \frac{dm_i \mathbf{v}_i}{dt}\right).$$

Hence every Lorentz observer sees that

$$\frac{dm_i}{dt} = -\sum_{\substack{i=1\\i\neq j}}^{n} \frac{dm_j}{dt},$$

$$\frac{dm_i\mathbf{v}_i}{dt} = -\sum_{\substack{i=1\\i\neq j}}^{n} \frac{dm_j\mathbf{v}_j}{dt} = \mathbf{f}_i.$$

Defining $\mathbf{f}_{ij} \triangleq -\dfrac{dm_j\mathbf{v}_j}{dt}$ as the interaction the j^{th} particle has on the i^{th} particle and defining $\mathbf{f}_i \triangleq \sum_{\substack{i=1\\i\neq j}}^{n} \mathbf{f}_{ij}$ we have that in every Lorentz coordinate system

$$\frac{dm_i\mathbf{v}_i}{dt} = \mathbf{f}_i,$$

where $\mathbf{f}_i$ is interpreted as the *classical force* acting on the particle. Thus Newton's second and third law holds in every Lorentz coordinate system. Using the Lorentz transformations one finds that in general $\dfrac{dm_i\mathbf{v}_i}{dt} \neq \dfrac{dm_i'\mathbf{v}_i'}{dt'}$ and thus the classical force acting observed in any two Lorentz coordinates is not the same unlike in Galilean mechanics.

Consider a particular Lorentz coordinate system $\mathbf{e}$. From the principle of conservation of momentum, analogous to Galilean mechanics we would be prompted to define the *force 4-vector F* by the relationship

$$F = \frac{dP}{d\tau} = \gamma\left(\frac{dm}{dt}, \frac{dm\mathbf{v}}{dt}\right) = \gamma\left(\frac{dm}{dt}, \mathbf{f}\right),$$

where $\mathbf{f}$ is the classical force acting on the particle as observed by $\mathbf{e}$.

Since $<< P,P >>= m_r^2$ we have that

$$<< \frac{dP}{d\tau}, P >>=<< F,P >>= m\gamma\frac{dm}{dt} - \frac{m}{c^2}\gamma\mathbf{f}\cdot\mathbf{v} = 0.$$

Which gives that

$$\frac{dm}{dt} = \frac{1}{c^2}\mathbf{f}\cdot\mathbf{v},$$

and hence that the force 4-vector is $F = \gamma\left(\frac{1}{c^2}\mathbf{f}\cdot\mathbf{v}, \mathbf{f}\right)$. Unlike in the case of Galilean mechanics we find that this quantity is not observer independent.

However from $\frac{dm}{dt} = \frac{1}{c^2}\mathbf{f}\cdot\mathbf{v}$ we see that all observers agree that mc^2 is equal to the energy imparted to the particle by the classical force $\mathbf{f}$.

Thus, in each Lorentz coordinate system, we may associate with a mass m the energy E by

$$E = mc^2,$$

in an coordinate independent manner. Note that one can re-arrange the right hand side to also read

$$E^2 = m_r^2 c^4 + c^2 ||\mathbf{p}||^2,$$

Notice that this is a direct consequence of Axiom-1.3. Further more this unites classical momentum and energy into the single quantity called the momentum 4-vector given by

$$P = \left(\frac{E}{c^2}, m\mathbf{v} \right)$$

Chapter 2
Quantum Mechanics

2.1 Introduction

Let M be a smooth manifold. Consider the Hilbert space $\mathscr{H}$ of complex valued smooth functions with compact support on M and a suitable innerproduct. A quantum mechanical system in the absence of an electromagnetic field is described by the Hilbert space $\mathscr{H}$ and a Hermitian operator H (the Hamiltonian) on $\mathscr{H}$. The first postulate of quantum mechanics is that the state of the system is described by an element $\psi(t,p) \in \mathscr{H}$ and evolves according to the Schrödinger equation

$$i\hbar\frac{\partial \psi}{\partial t} = H\psi. \tag{2.1}$$

The magnitude $|\psi(t,p)|^2$ is assumed to give the probability of finding the system at $p \in M$ at time t.

The second postulate of quantum mechanics state that measurements correspond to Hermitian operators and the only possible values that can be obtained in a measurement of the system must correspond to an eigenvalue of the Hermitian operator that describes the observable. Thus if B is a Hermitian operator associated with a certain observable of the systems. Upon a measurement of the observable, corresponding to B, the only possible values that it can take are the eigenvalues of B. The expected value of B, when the system is in the state ψ is

$$\langle B\rangle = \langle \psi|B\psi\rangle = \langle \psi|B|\psi\rangle.$$

Since B is Hermitian this value is always real. The time evolution of the expected value of the observable B is given by [1]

$$i\hbar\frac{d\langle B\rangle}{dt} = \langle \psi\,|[B,H]|\,\psi\rangle + \left\langle \psi\left|\frac{\partial B}{\partial t}\right|\psi\right\rangle = \langle [B,H]\rangle + i\hbar\left\langle\frac{\partial B}{\partial t}\right\rangle. \tag{2.2}$$

We restrict attention to operators that do not explicitly depend on time. Then the above equation becomes

$$i\hbar\frac{d\langle B\rangle}{dt} = \langle [B,H]\rangle. \tag{2.3}$$

[1] For two Hermitian operators A and B it can be shown that $\langle [A,B]\rangle$ is always imaginary and $\langle AB+BA\rangle$ is always real.

The standard deviation (uncertainty) of the observable B is given by the square root of

$$S_B = \langle B^2\rangle - \langle B\rangle^2 = \langle (B - \langle B\rangle)^2\rangle.$$

We can also write the time evolution of the variance (the standard deviation squared) as

$$i\hbar\frac{dS_B}{dt} = \left\langle \left[B^2 - 2\langle b\rangle B, H_0\right]\right\rangle = \langle [S_B, H_0]\rangle\,. \tag{2.4}$$

2.2 First order Perturbations and Scattering

The formal solution to the unperturbed Schrödinger equation (2.1) with Hamiltonian H_0 is given by

$$\psi_0(t) = e^{\frac{-itH_0}{\hbar}}\psi_0(0).$$

When $\psi_0(0) = \phi_n$

$$\psi_0(t) = e^{\frac{-itH_0}{\hbar}}\phi_n = e^{\frac{-iE_nt}{\hbar}}\phi_n$$

and the state is called a pure state. Define

$$U_0(t) \triangleq e^{\frac{-iH_0t}{\hbar}}.$$

When $H^c \neq 0$ we seek to write the solution to the Schrödinger equation, $\psi(t)$, with perturbed Hamiltonian $H = H_0 + H_c$ as

$$\psi(t) = U_0(t)\varphi(t)$$

for some wave function $\varphi(t)$. Differentiating this we find that $\varphi(t)$ must then satisfy

$$i\hbar\frac{\partial\varphi(t)}{\partial t} = (U_0^\dagger H^c U_0)\varphi(t) = \mathbf{H}^c(t)\varphi(t),$$

where we have defined $\mathbf{H}^c(t) \triangleq (U_0^\dagger H^c U_0)$. From which we find

$$\varphi(t) = \varphi(0) + \frac{1}{i\hbar}\int_0^t \mathbf{H}^c(t')\varphi(t')dt'.$$

This integral can be evaluated iteratively as

$$\varphi_{n+1}(t) = \varphi(0) + \frac{1}{i\hbar}\int_0^t \mathbf{H}^c(t')\varphi_n(t')dt',$$

$$\varphi(t) = \varphi(0) + \sum_{n=1}^{\infty}\frac{1}{(i\hbar)^n}\int_0^t dt_1\int_0^{t_1} dt_2\cdots\int_0^{t_{n-1}}(\mathbf{H}^c(t_1)\mathbf{H}^c(t_2)\cdots\mathbf{H}^c(t_n))\varphi(0)dt_n.$$

This can be written as

$$\varphi(t) = \varphi(0) + \sum_{n=1}^{\infty}\frac{1}{n!\,(i\hbar)^n}\int_0^t\int_0^t\cdots\int_0^t \mathscr{T}(\mathbf{H}^c(t_1)\mathbf{H}^c(t_2)\cdots\mathbf{H}^c(t_n))\varphi(0)dt_1dt_2\cdots dt_n,$$

$$\triangleq \mathscr{T} \exp\left(\frac{1}{i\hbar}\int_0^t \mathbf{H}^c(\tau)d\tau\right)\varphi(0).$$

where the chronological operator $\mathscr{T}$ orders the terms with time is increasing from right to left. This is called the Dyson series.

Using the above Dyson series when H_c is small we see that

$$\varphi(t) \approx \varphi(0) + \frac{1}{i\hbar}\int_0^t \mathbf{H}^c(t')\varphi(0)dt',$$

and hence since $\psi(0) = \varphi(0)$ and $\psi_0(t) = U_0(t)\psi(0)$ we have that the first order approximation of the solution is that

$$\psi(t) \approx \psi_0(t) + \frac{1}{i\hbar}\int_0^t e^{\frac{-i(t-t')H_0}{\hbar}} H^c \psi_0(t')dt',$$

Thus if $\psi(0) = \phi_n$ one of the eigenstates then

$$\psi(t) \approx U_0(t)\phi_n + \frac{1}{i\hbar}U_0(t)\int_0^t \mathbf{H}^c\phi_n dt' = U_0(t)\left(\phi_n + \frac{1}{i\hbar}\int_0^t \mathbf{H}^c\phi_n dt'\right).$$

Note that we can write

$$\psi(t) = \sum_k c_{n\to k}U_0(t)\phi_k = \sum_k c_{n\to k}e^{\frac{-iE_k t}{\hbar}}\phi_k.$$

Here $P_{n\to k} = |c_{n\to k}|^2$ is interpreted to be the transition probability from the state ϕ_n to ϕ_k within the time t, and is given by

$$c_{n\to k} = \langle U_0(t)\phi_k, \psi(t)\rangle = \delta_{k,n} + \frac{1}{i\hbar}\int_0^t \phi_k^\dagger(U_0^\dagger(t')H^c U_0(t'))\phi_n dt' = \delta_{k,n} + \frac{1}{i\hbar}\int_0^t e^{-i\omega_{nk}t'}\phi_k^\dagger H^c\phi_n dt'$$

where $\hbar\omega_{nk} = E_n - E_k$. Sice the probability of the state ϕ_n decaying to ϕ_k in time t is $P_{n\to k}$, the time it takes for ϕ_n to decay into ϕ_k with probability one is

$$\tau_{n\to k} = \frac{t}{P_{n\to k}}.$$

This is called the relaxation time of the interaction.

2.2.1 H^c is time independent

$$c_{n\to k} = \delta_{k,n} + \frac{1}{\hbar\omega_{nk}}\left(e^{-i\omega_{nk}t} - 1\right)\langle\phi_k|H^c|\phi_n\rangle$$

For $k \neq n$

$$P_{n\to k} = \frac{1}{\hbar^2}\frac{\sin^2\left(\frac{\omega_{nk}t}{2}\right)}{\left(\frac{\omega_{nk}}{2}\right)^2}|\langle\phi_k|H^c|\phi_n\rangle|^2 = \frac{t^2}{\hbar^2}\mathrm{sinc}^2\left(\frac{\omega_{nk}t}{2}\right)|\langle\phi_k|H^c|\phi_n\rangle|^2$$

It can be shown that
$$\lim t \to \infty \frac{\sin^2\left(\frac{\omega_{nk}t}{2}\right)}{\left(\frac{\omega_{nk}}{2}\right)^2} = 2\pi t\,\delta(\omega_{nk}).$$
Thus
$$P_{n\to k} = 2\pi t\,\frac{1}{\hbar^2}|\langle\phi_k|H^c|\phi_n\rangle|^2\delta(\omega_{nk}).$$
Thus for sufficiently large t the transition probability is not zero only if $(E_n - E_k)/\hbar = \omega_{nk} = 0$ (energy conservation). This means that if the unperturbed Hamiltonian is nondegenerate (that is $E_k \neq E_n$ for all $k \neq n$) then spontaneous decay does not occur. The relaxation time of the interaction is given by
$$\frac{1}{\tau_{n\to k}} = \frac{2\pi}{\hbar^2}|\langle\phi_k|H^c|\phi_n\rangle|^2\delta(\omega_{nk}).$$

2.2.2 H^c is periodic in time

Let $H^c = \bar{H}^c\cos\omega t = \frac{\bar{H}^c}{2}(e^{-i\omega t} + e^{i\omega t})$.
$$\begin{aligned} c_{n\to k} &= \delta_{n,k} + \frac{1}{i\hbar}\int_0^t \phi_k^\dagger(U_0^\dagger H^c U_0)\phi_n dt' = \delta_{n,k} + \frac{1}{i\hbar}\int_0^t e^{-i\omega_{nk}t'}\phi_k^\dagger H^c\phi_n dt' \\ &= \delta_{k,n} + \frac{\langle\phi_k|\bar{H}^c|\phi_n\rangle}{2\hbar}\left(\frac{\left(e^{-i(\omega+\omega_{nk})t} - 1\right)}{(\omega+\omega_{nk})} + \frac{\left(e^{-i(\omega_{nk}-\omega)t} - 1\right)}{(\omega_{nk}-\omega)}\right) \end{aligned}$$

Then for $n \neq k$
$$P_{n\to k} = \frac{1}{4\hbar^2}\left(\frac{\sin^2\left(\frac{(\omega-\omega_{nk})t}{2}\right)}{\left(\frac{(\omega-\omega_{nk})}{2}\right)^2} + \frac{\sin^2\left(\frac{(\omega+\omega_{nk})t}{2}\right)}{\left(\frac{(\omega+\omega_{nk})}{2}\right)^2}\right)|\langle\phi_k|\bar{H}^c|\phi_n\rangle|^2$$
and for sufficiently large times
$$P_{n\to k} = \frac{\pi t}{2\hbar^2}|\langle\phi_k|\bar{H}^c|\phi_n\rangle|^2(\delta(\omega-\omega_{nk}) + \delta(\omega+\omega_{nk})).$$

Thus for sufficiently large t the transition probability is not zero only if
$$E_n - E_k = \hbar\omega,$$
or
$$E_k - E_n = \hbar\omega.$$
For $E_n - E_k = \hbar\omega$ the transition probability is
$$P_{n\to k} = \frac{\pi t}{2\hbar^2}|\langle\phi_k|\bar{H}^c|\phi_n\rangle|^2.$$
The relaxation time of the interaction is given by
$$\frac{1}{\tau_{n\to k}} = \frac{\pi}{2\hbar^2}|\langle\phi_k|\bar{H}^c|\phi_n\rangle|^2.$$

2.3 Quantum Particle

$$H_0 = -\frac{\hbar^2}{2}\frac{\partial^2}{\partial z^2} + V(z,t) \tag{2.5}$$

Let $\widehat{z}$ be the position operator and $\widehat{p}$ be the momentum operator given by

$$\widehat{p} = -i\hbar\frac{\partial}{\partial z}, \qquad \widehat{z} = z.$$

It can be shown that

$$[\widehat{z}, H_0] = i\hbar\widehat{p}, \qquad [\widehat{p}, H_0] = -i\hbar\frac{\partial V}{\partial z}, \qquad [\widehat{p}, \widehat{z}] = -i\hbar$$

From (**??**) we have

$$\frac{dz}{dt} = p - \frac{i}{\hbar}\langle[\widehat{z}, H^c]\rangle, \tag{2.6}$$

$$\frac{dp}{dt} = -\left\langle \frac{\partial V}{\partial z} \right\rangle - \frac{i}{\hbar}\langle[\widehat{p}, H^c]\rangle, \tag{2.7}$$

where $z = \langle\widehat{z}\rangle$, $p = \langle\widehat{p}\rangle$ and $h_0 = \langle H_0\rangle$. From (2.4) we have that

$$\frac{dS_z}{dt} = \langle\widehat{z}\widehat{p} + \widehat{p}\widehat{z}\rangle - 2zp - \frac{i}{\hbar}\langle[S_z, H^c]\rangle \tag{2.8}$$

$$\frac{dS_p}{dt} = -\left\langle \widehat{p}\frac{\partial V}{\partial z} + \frac{\partial V}{\partial z}\widehat{p} \right\rangle + 2\left\langle \frac{\partial V}{\partial z} \right\rangle p - \frac{i}{\hbar}\langle[S_p, H^c]\rangle \tag{2.9}$$

Choosing $H_c = -u(t)\widehat{z}$ corresponds to a quantum particle moving under the influence of an external force $u(t)$. For instance this could model a charged molecule in a electric filed of strength $u(t)$. Then we have

$$\frac{dz}{dt} = p, \tag{2.10}$$

$$\frac{dp}{dt} = -\left\langle \frac{\partial V}{\partial z} \right\rangle + u(t), \tag{2.11}$$

$$\frac{dS_z}{dt} = \langle\widehat{z}\widehat{p} + \widehat{p}\widehat{z}\rangle - 2zp \tag{2.12}$$

$$\frac{dS_p}{dt} = -\left\langle \frac{\partial V}{\partial z}\widehat{p} + \widehat{p}\frac{\partial V}{\partial z} \right\rangle + 2\left\langle \frac{\partial V}{\partial z} \right\rangle p. \tag{2.13}$$

Thus the expected values of a quantum mechanical particle behaves like the dynamics of a classical particle.

The controls do not appear in the variance equation. Thus the challenge is to see if we can drive the expected values desirably while ensuring that the uncertainties remain bounded. Or may be if we are only interested in driving the expected value of the position we need to only ensure that variance in the expected value stays reasonably bounded while we would not be concerned of the uncertainty in the momentum.

A thing of practical importance is to ensure that the Heisenberg uncertainty principle is always satisfied. That is

$$S_z S_p \geq \frac{\hbar^2}{4}$$

If we can prove this we are not violating any physics.

From (2.12) and (2.13) we have when $V(z) = \omega^2 z^2/2$ that

$$\frac{d}{dt}(S_p + \omega^2 S_z) = 0. \tag{2.14}$$

Which implies that $(S_p + \omega^2 S_z) = \gamma$, a constant. Differentiating (2.14) we have that

$$\frac{dS_z S_p}{dt} = f(t)(S_p - S_z\omega^2)$$

....

Now if the expected values of the operators are measurable then the problem becomes trivial as we can set $u(t) = f(z,p)$ such that the expected values are driven as desirably. If not we are stuck with open-loop control. Unfortunately this is all known to physicists. The quantum control guys should know this but do not seem to mention it. This is interesting if we can measure expected values. This may not be possible after all. The act of measuring introduces another interacting Hamiltonian H^o. This may be modeled as disturbances acting on the expected value system. Letting $d_1(t) = \langle [\widehat{z}, H^o] \rangle$ and $d_2(t) = \langle [\widehat{p}, H^o] \rangle$ we then have

$$\frac{dz}{dt} = p + d_1(t), \tag{2.15}$$

$$\frac{dp}{dt} = -\left\langle \frac{\partial V}{\partial z} \right\rangle + d_2(t) + u(t). \tag{2.16}$$

2.3.1 Quantum Harmonic Oscillator

$$H_0 = -\frac{\hbar^2}{2}\frac{\partial^2}{\partial z^2} + \frac{\omega^2}{2} z^2 \tag{2.17}$$

Define

$$\widehat{a}^\dagger = \frac{1}{\sqrt{2\hbar\omega}}\left(\omega z - \hbar\frac{\partial}{\partial z}\right), \quad \widehat{a} = \frac{1}{\sqrt{2\hbar\omega}}\left(\omega z + \hbar\frac{\partial}{\partial z}\right)$$

From these it also follows that

$$\widehat{z} = \sqrt{\frac{\hbar}{2\omega}}\left(\widehat{a}^\dagger + \widehat{a}\right), \quad \widehat{p} = -i\hbar\frac{\partial}{\partial z} = i\sqrt{\frac{\hbar\omega}{2}}\left(\widehat{a}^\dagger - \widehat{a}\right)$$

It can be shown that

$$H_0 = \hbar\omega\left(\frac{1}{2} + \widehat{a}^\dagger\widehat{a}\right) = \hbar\omega\left(-\frac{1}{2} + \widehat{a}\widehat{a}^\dagger\right) \tag{2.18}$$

The eigenvectors of H_0 are given by

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}}\,\pi^{-1/4}\exp(-x^2/2)\,H_n(x).$$

where $H_n(x)$ are the Hermite polynomials and the corresponding eigenvalues are given by

$$E_n = \hbar\omega\left(\frac{1}{2}+n\right).$$

It can be shown that

$$\widehat{a}^{\dagger}\widehat{a}\psi_n = n\,\psi_n$$

and

$$\widehat{a}^{\dagger}\psi_n = \sqrt{n+1}\,\psi_{n+1},\;\;\widehat{a}\psi_n = \sqrt{n}\,\psi_{n-1}$$

In addition the raising and lowering operators satisfy the conditions

$$[\widehat{a}^{\dagger},\widehat{a}] = -1,\;\;[\widehat{a}^{\dagger}\widehat{a},\widehat{a}^{\dagger}] = \widehat{a}^{\dagger},\;\;[\widehat{a}^{\dagger}\widehat{a},\widehat{a}] = -\widehat{a}.$$

$$[\widehat{z},H_0] = i\hbar\widehat{p},\;\;[\widehat{p},H_0] = -i\hbar\omega^2\widehat{z}\;\;[\widehat{p},\widehat{z}] = -i\hbar$$

2.3.1.1 Two coupled Harmonic Oscillators

$$H = H_0 + H_I$$

$$H_0 = H_1 + H_2$$

$$H_i = -\frac{\hbar^2}{2}\frac{\partial^2}{\partial z_i^2} + \frac{\omega_i^2}{2}z_i^2 = \hbar\omega_i\left(a_i^{\dagger}a_i + \frac{1}{2}\right)$$

The wave function of the system is $\psi(z_1,z_2)$. Let $N_i = a_i^{\dagger}a_i$. We also have

$$[a_i^{\dagger},a_j] = -\delta_{ij}.$$

Let $\phi_{i_n}(z_i)$ be eigenfunctions of H_i. That is

$$H_i\phi_{i_n} = E_{in}\phi_{i_n}.$$

Then

$$H_0\psi_{nm} = E_{nm}\psi_{nm}$$

where $\psi_{nm}(z_1,z_2) = \phi_{1n}(z_1)\phi_{2m}(z_2)$ and $E_{nm} = E_{1n}+E_{2m}$. The $\psi_{nm}(z_1,z_2)$ form a basis for the Hilbert space of complex valued functions of two variables. This is in general true for any Hermitian operator A_i That is

$$A_i\phi_{i_n} = \lambda_{in}\phi_{i_n}.$$

Then

$$(A_1+A_2)\psi_{nm} = \lambda_{nm}\psi_{nm}$$

where $\psi_{nm}(z_1,z_2) = \phi_{1n}(z_1)\phi_{2m}(z_2)$ and $\lambda_{nm} = \lambda_{1n}+\lambda_{2m}$ and the $\psi_{nm}(z_1,z_2)$ form a basis.

Let ψ be a general state of the interacting system. Now it may no longer be true that it can be written as

$$\psi = \phi_1 \phi_2.$$

However it can be written in the form

$$\psi(z_1, z_2) = \sum_{nm} c_{nm} \psi_{nm}.$$

We know

$$\phi_{in}(z_i, t) = e^{\frac{-iH_i t}{\hbar}} \phi_{in}(z_i) = e^{\frac{-iE_{in} t}{\hbar}} \phi_{in}(z_i)$$

We want to write

$$\psi(z_1, z_2, t) = e^{\frac{-iHt}{\hbar}} \phi_{1n}(z_1) \phi_{2m}(z_2)$$

If $H_I = 0$ then

$$\psi_0(z_1, z_2, t) = e^{\frac{-iH_0 t}{\hbar}} \phi_{1n}(z_1) \phi_{2m}(z_2) = e^{\frac{-iE_{1n} t}{\hbar}} \phi_{1n}(z_1) e^{\frac{-iE_{2m} t}{\hbar}} \phi_{2m}(z_2) = e^{\frac{-i(E_{1n}+E_{2m})t}{\hbar}} \phi_{1n}(z_1) \phi_{2m}(z_2).$$

and the state is a pure state.

When $H_I \neq 0$

$$\psi(z_1, z_2, t) = e^{\frac{-iHt}{\hbar}} \phi_{1n}(z_1) \phi_{2m}(z_2) = e^{\frac{-iHt}{\hbar}} \psi_{nm}(z_1, z_2) = \sum_{jk} c_{jk} e^{\frac{-iH_0 t}{\hbar}} \psi_{jk}$$

Since we know what $U_0(t) = e^{\frac{-iH_0 t}{\hbar}}$ is we want to be able to write $\psi(z_1, z_2, t)$ as

$$\psi(z_1, z_2, t) = U_0(t) \varphi(z_1, z_2, t)$$

Differentiating this we find that $\varphi(z_1, z_2, t)$ satisfies

$$i\hbar \varphi(z_1, z_2, t) = (U_0^\dagger H_I U_0) \varphi(z_1, z_2, t)$$

From which we find

$$\varphi(z_1, z_2, t) = \varphi(z_1, z_2, 0) - \frac{1}{i\hbar} \int_0^t (U_0^\dagger H_I U_0) \varphi(z_1, z_2, t') dt'$$

Using the Dyson series this can be approximated as

$$\varphi(z_1, z_2, t) \approx \varphi(z_1, z_2, 0) - \int_0^t (U_0^\dagger H_I U_0) \varphi(z_1, z_2, 0) dt'$$

Thus

$$\psi(z_1, z_2, t) \approx U_0 \psi_{nm} - \frac{1}{i\hbar} U_0 \int_0^t (U_0^\dagger H_I U_0) \psi_{nm} dt'$$

We can write

$$\psi(z_1, z_2, t) = \sum_{jk} c_{nm \to jk} U_0 \psi_{jk} = \sum_{jk} c_{nm \to jk} e^{\frac{iE_{jk}}{\hbar}} \psi_{jk}$$

where we interpret $P_{nm\to jk} = |c_{nm\to jk}|^2$ to be the transition probability from the product state ψ_{nm} to ψ_{jk} and is given by

$$c_{nm\to jk} = \delta_{jk,nm} - \frac{1}{i\hbar}\int_0^t \psi_{jk}^\dagger (U_0^\dagger H_I U_0)\psi_{nm} dt' = \delta_{jk,nm} - \frac{1}{i\hbar}\int_0^t e^{\frac{-i(E_{nm}-E_{jk})t}{\hbar}} \psi_{jk}^\dagger H_I \psi_{nm} dt'$$

If H_I is time independent

$$c_{nm\to jk} = \delta_{jk,nm} - \frac{1}{(E_{nm}-E_{jk})}\left(e^{\frac{-i(E_{nm}-E_{jk})t}{\hbar}} - 1\right)\langle \psi_{jk}|H_I|\psi_{nm}\rangle$$

For $jk \neq nm$

$$P_{nm\to jk} = \frac{1}{\hbar^2}\frac{\sin^2\left(\frac{(\omega_{nm}-\omega_{jk})t}{2}\right)}{\left(\frac{(\omega_{nm}-\omega_{jk})}{2}\right)^2}|\langle \psi_{jk}|H_I|\psi_{nm}\rangle|^2$$

It can be shown that

$$\lim t \to \infty \frac{\sin^2\left(\frac{(\omega_{nm}-\omega_{jk})t}{2}\right)}{\left(\frac{(\omega_{nm}-\omega_{jk})}{2}\right)^2} = 2\pi t\,\delta(\omega_{nm}-\omega_{jk}).$$

Thus for large t the transition probability is nonzero only if $\omega_{nm} = \omega_{jk}$ (energy conservation). For sufficiently large t we have

$$P_{nm\to jk} = \frac{2\pi t}{\hbar^2}|\langle \psi_{jk}|H_I|\psi_{nm}\rangle|^2\delta(\omega_{nm}-\omega_{jk})$$

$$\omega_{nm}-\omega_{jk} = \omega_1\left(\frac{1}{2}+n\right)+\omega_2\left(\frac{1}{2}+m\right)-\omega_1\left(\frac{1}{2}+j\right)-\omega_2\left(\frac{1}{2}+k\right)$$

$$\omega_{nm}-\omega_{jk} = \omega_1(n-j)+\omega_2(m-k)$$

We seek a 2-phonon interaction where one phonon of mode 1 decays into one phonon of mode two. Thus we seek $j = n-1, k = m+1$. A transition of this nature will occur with non zero probability only if

$$\omega_{nm} - \omega_{(n-1)(m+1)} = \omega_1 - \omega_2 = 0.$$

Observe that for the transition $j = n+1, k = m-1$ we require $\omega_1 - \omega_2 = 0$. Thus having $\omega_1 = \omega_2$ will yield a net transition rate (transition probability per unit time) for one phonon of mode 1 decaying to one phonon of mode 2 is given by

$$W_{2\to 1} = \frac{2\pi}{\hbar^2}\left(|\langle \psi_{(n-1)(m+1)}|H_I|\psi_{nm}\rangle|^2 - |\langle \psi_{(n+1)(m-1)}|H_I|\psi_{nm}\rangle|^2\right).$$

We are interested in the case where $\omega_1 > \omega_2$, then clearly $\omega_1 \neq \omega_2$ then no such transition will occur. However if we select $H_I = \bar{H}_I \cos\omega t$ then from time varying perturbation theory

$$W_{2\to 1} = \frac{2\pi}{4\hbar^2}\left(|\langle \psi_{(n-1)(m+1)}|\bar{H}_I|\psi_{nm}\rangle|^2 - |\langle \psi_{(n+1)(m-1)}|\bar{H}_I|\psi_{nm}\rangle|^2\right)\left(\delta(\omega_{nm}-\omega_{jk}-\omega)+\delta(\omega_{nm}-\omega_{jk}+\omega)\right)$$

Thus now a three phonon interaction is possible if

$$\omega_{nm} - \omega_{(n-1)(m+1)} - \omega = \omega_1 - \omega_2 - \omega = 0.$$

or

$$\omega_{nm} - \omega_{(n-1)(m+1)} + \omega = \omega_1 - \omega_2 + \omega = 0.$$

Since $\omega_1 > \omega_2$ this transition can be induced if we let the forcing frequency to be

$$\omega = \omega_1 - \omega_2.$$

and the decay rate of mode one phonons is

$$W = \frac{\pi}{2\hbar^2}\left(|\langle\psi_{(n-1)(m+1)}|\bar{H}_I|\psi_{nm}\rangle|^2 - |\langle\psi_{(n+1)(m-1)}|\bar{H}_I|\psi_{nm}\rangle|^2\right).$$

....

When $H_I = a_2^\dagger a_1^2$,

$$\langle\psi_{jk}|a_2^\dagger a_1^2|\psi_{nm}\rangle = \sqrt{n(n-1)(m+1)}\,\delta_{jk,(n-2)(m+1)},$$

and is non zero only when $j=(n-2), k=(m+1)$. When $H_I = a_2 a_1^{\dagger 2}$,

$$\langle\psi_{jk}|a_2 a_1^{\dagger 2}|\psi_{nm}\rangle = \sqrt{m(n+1)(n+2)}\,\delta_{jk,(n+2)(m-1)},$$

and is not zero only if $j=(n+2), k=(m-1)$.

Similarly if we let $H_I = a_1^\dagger a_2^2$. Then

$$\langle\psi_{jk}|a_1^\dagger a_2^2|\psi_{nm}\rangle = \sqrt{m(m-1)(n+1)}\,\delta_{jk,(n+1)(m-2)}$$

and is not zero only if $j=(n+1), k=(m-2)$. Let $H_I = a_1 a_2^{\dagger 2}$. Then

$$\langle\psi_{jk}|a_1 a_2^{\dagger 2}|\psi_{nm}\rangle = \sqrt{n(m+1)(m+2)}\,\delta_{jk,(n-1)(m+2)}$$

and is not zero only if $j=(n-1), k=(m+2)$. It can be shown that other cubic terms give $\langle\psi_{(n-2)(m+1)}|H_I|\psi_{nm}\rangle = \langle\psi_{(n+2)(m-1)}|H_I|\psi_{nm}\rangle = 0$. Thus if H_I is any cubic nonlinearity in z_1, z_2, only the term $z_2 z_1^2$ contributes to a nonzero $\langle\psi_{jk}|H_I|\psi_{nm}\rangle$. Thus if the coefficient of the term $z_2 z_1^2$ is γ then

$$W_{2\to 1} = \frac{\gamma\pi}{4}\sqrt{\frac{1}{2\hbar\omega_1^2\omega_2}}\left(n(n-1)(m+1) - m(n+1)(n+2)\right).$$

Chapter 3
Maxwell's Laws of Electromagnetism

Consider an inertial observer **e** and let the associated Lorentz coordinate system for space-time be $(t, \mathbf{x}) = (t, x_1, x_2, x_3)$. We assume that space if filled with particles with a certain density. For a given observer **e** some of these particles will appear stationary while for some of them appear to move. We have already assumed that each particle has an inertial observer invariant quantity called the rest mass m_r. Similarly we will also assume the following:

Axiom 3.1 Charge of a particle: *Every particle has a property called rest charge that is observed to be the same for all inertial observers.*

Denote by ρ the density of the stationary distribution of the charged particles and by ρ^J the current density, per unit area, of the flow of particles at a given spatial point P as observed by the inertial observer **e**. Clearly different inertial observers will not see that these quantities are the same. We consider the problem of estimating the motion of a charged particle P of charge q that is moving at a velocity $\mathbf{v}$ as observed in **e**. It was discovered that this motion is described by the Maxwell's equations.

Let us begin by defining the spatial forms

$$\begin{aligned}
\mathscr{E}_s &= E_1 dx^1 + E_2 dx^2 + E_3 dx^3, \\
\mathscr{B}_s &= B_1 dx^2 \wedge dx^3 + B_2 dx^3 \wedge dx^1 + B_3 dx^1 \wedge dx^2, \\
\mathscr{J}_s &= J_1 dx^2 \wedge dx^3 + J_2 dx^3 \wedge dx^1 + J_3 dx^1 \wedge dx^2, \\
\mathscr{D}_s &= \rho dx^1 \wedge dx^2 \wedge dx^3
\end{aligned}$$

where $\mathscr{E}_s$ is the electric field 1-form associated with $\mathbf{E} = (E_1, E_2, E_3)$, $\mathscr{B}_s$ is the magnetic field 2-form associated with $\mathbf{B} = (B_1, B_2, B_3)$, $\mathscr{J}_s$ is the current 2-form associated with $\mathbf{J}_s = (J_1, J_2, J_3)$, and $\mathscr{D}_s$ is the charge density.

Associated with these spatial quantities one defines the following true forms on 4-D Minkowski space as follows:

$$\mathscr{F} = \mathscr{E}_s \wedge dt + \frac{1}{c}\mathscr{B}_s, \tag{3.1}$$

$$\mathscr{J} = \mathscr{D}_s - \mathscr{J}_s \wedge dt, \tag{3.2}$$

where

Recall that we require that Laws of nature be Lorentz invariant. That is, the mathematical expressions describing the laws of nature to be the same in all Lorentz coordinate systems. It is clear that the spatial quantities such as $\mathbf{v}$, $\mathbf{f}$, $\mathbf{J}$, ρ, are not Lorentz invariant the quantities where as V, F, $\mathscr{J}$ are Lorentz invariant by virtue of being intrinsic objects. Thus we see that Laws of nature should be expressed using such intrinsic quantities if they are to be Lorentz invariant and not using $\mathbf{v}$, $\mathbf{f}$, $\mathbf{J}$, ρ. In Axiom-1.3 we saw the first of such laws which stated that the magnitude of the 4-momentum of any particle is a Lorentz invariant quantity called the rest mass. Then in Axiom-1.4 it was stated that the total 4-momentum of an interacting but isolated set of particles remained constant in any Lorentz coordinate system. These two axioms combined implied that every inertial observer sees the energy mass relationship $E = mc^2$.

Summarising a multitude of experimental observations, James Klerk Maxwell along with Hendrik Lorentz observed that the law governing the motion of charged particles can be stated as follows:

Axiom 3.2 Maxwell's Laws of Electromagnetism: *In every Lorentz coordinate system (ie. for every inertial observer) the force acting on a particle of charge q moving at a speed of $\mathbf{v}$ with respect to the Lorentz coordinates is the Lorentz force 4-vector F that is explicitly given by*

$$F^\flat = \frac{1}{c^2} q\, i_U \mathscr{F}, \tag{3.3}$$

where $U = \gamma(1, \mathbf{v})$ is the velocity 4-vector of the charged particle with respect to the proper time parameter τ, and $\mathscr{F}$ is the unique closed 2-form such that

$$d * \mathscr{F} = 4\pi \mathscr{J}. \tag{3.4}$$

Notice that any 2-form in Minkoswski space can be written as (3.1). Since $\mathscr{F}$ is closed $d\mathscr{F} = 0$ and thus

$$d\mathscr{F} = \left(\mathbf{d}_s \mathscr{E}_s + \frac{1}{c}\frac{\partial \mathscr{B}_s}{\partial t}\right) \wedge dt + \frac{1}{c}\mathbf{d}_s \mathscr{B}_s = 0,$$

yields the first two Maxwell's equations. Also recall that

$$\begin{aligned} *\mathscr{F} &= \star\mathscr{E}_s - c\star\mathscr{B}_s \wedge dt, \\ d * \mathscr{F} &= \mathbf{d}_s \star \mathscr{E}_s + \left(\frac{\partial \star \mathscr{E}_s}{\partial t} - c\,\mathbf{d}_s \star \mathscr{B}_s\right) \wedge dt, \end{aligned}$$

and hence that $d * \mathscr{F} = 4\pi \mathscr{J}$ corresponds to the last two Maxwell's equations. Thus, in summary, the assumptions $d\mathscr{F} = 0$ and $d * \mathscr{F} = 4\pi \mathscr{J}$ result in the four Maxwell's equations,

$$\mathbf{d}_s \mathscr{E}_s = -\frac{1}{c}\frac{\partial \mathscr{B}_s}{\partial t}, \tag{3.5}$$

$$\mathbf{d}_s \mathscr{B}_s = 0, \tag{3.6}$$

$$\mathbf{d}_s \star \mathscr{E}_s = 4\pi \mathscr{D}_s, \tag{3.7}$$

$$\mathbf{d}_s \star \mathscr{B}_s = \frac{1}{c}\frac{\partial \star \mathscr{E}_s}{\partial t} + \frac{4\pi}{c}\mathscr{J}_s. \tag{3.8}$$

Conversely we see that the Maxwell's equations in the Lorentz co-ordinates $\mathbf{e}$ implies that every inertial observer sees that $d\mathscr{F} = 0$ and $d * \mathscr{F} = 4\pi \mathscr{J}$. Equation (3.5) is called *Faraday's Law*, Equations (3.6) and (3.7) are *Gauss's Law* for magnetic fields and electric fields, and equation (3.8) is called the *Ampere-Maxwell Law*. From the expression for $d\mathscr{F}$ we see that if an inertial observer $\mathbf{e}$ sees that Gauss's law of $\mathbf{d}_s \mathscr{B}_s = 0$ holds then $\mathbf{e}$ sees that there exists a 4-vector W with no time component such that $i_W \mathrm{vol} = d\mathscr{F} \neq 0$. If W' is the representation of W in a different inertial frame e', in general W' will have a time component. Thus e' will not observe $\mathbf{d}_s \mathscr{B}'_s = 0$. Thus if all inertial observers agree that $\mathbf{d}_s \mathscr{B}'_s = 0$ then necessarily $W = 0$ and hence $i_W \mathrm{vol} = d\mathscr{F} = 0$.

Thus if all inertial observers see that Gauss's law of absence of magnetic monopoles is true then necessarily they also see that Faraday's Law is true. A similar argument shows that if every inertial observers sees that Gauss's Law of electric fields is true then they also see that Ampere-Maxwell Law is true. This shows that Gauss's law implies that Faraday'a law and Ampere-Maxwell Law are true. Thus Gauss's law is a more fundamental law of nature.

Since $4\pi\, d\mathscr{J} = d\, d * \mathscr{F} = 0$, if D is a 4-dimensional compact space time region with boundary ∂D, then

$$\int_{\partial D} \mathscr{J} = \int_D d\mathscr{J} = 0.$$

This is a statement that *charge is conserved* in an space-time region. From

$$d\mathscr{J} = dt \wedge \frac{\partial \mathscr{J}}{\partial t} + \mathbf{d}_s \mathscr{J} = dt \wedge \left(\frac{\partial \mathscr{D}_s}{\partial t} + \mathbf{d}_s \mathscr{J}_s \right) = 0,$$

we obtain the equivalent version of conservation of charge given by

$$\frac{\partial \mathscr{D}_s}{\partial t} + \mathbf{d}_s \mathscr{J}_s = 0. \tag{3.9}$$

Let V^2 be a fixed 2-dimensional compact set in $\mathbb{R}^3$ with ∂V^2 being its boundary. Then the set $W^2(t) = (t, V^2(t))$ (a tubular region) is a 2-dimensional compact set in Minkowski space-time. Since when restricted to $W(t)$ the form dt is zero we see that

$$\int_{W^2(t)} \mathscr{F} = \frac{1}{c}\int_{V^2(t)} \mathscr{B}_s.$$

Thus from Section-7.4.1 we have seen that

$$\frac{d}{dt}\int_{V^2(t)} \mathscr{B}_s = \frac{d}{dt}\int_{W^2(t)} c\mathscr{F} = -\int_{\partial V^2(t)} c\left(\mathscr{E}_s - \frac{1}{c} i_{\mathbf{v}}\mathscr{B}_s\right). \tag{3.10}$$

This equation describes the electromotive force that if generated in a loop $\partial V(t)$ that is moving with velocity $\mathbf{v}(\mathbf{x})$ with respect to the inertial frame $e = (t, \mathbf{x})$.

Expressing the equations (3.5)–(3.8) using the coefficients of $\mathscr{E}_s$ and $\mathscr{B}_s$ given by $\mathbf{E}$ and $\mathbf{B}$ respectively we have the familiar form of the Maxwell's equations in cgs (Gaussian) units.

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \tag{3.11}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{3.12}$$

$$\nabla \times \mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c}\mathbf{J}, \tag{3.13}$$

$$\nabla \cdot \mathbf{E} = 4\pi\rho. \tag{3.14}$$

Given a smooth stationary charge distribution ρ and and a current density $\mathbf{J}$ in the $\mathbf{e}$ frame it is known that a unique solution $\mathbf{E}$ and $\mathbf{B}$ exists for the above set of PDEs. These quantities are called the *Electric field* and the *Magnetic field* observed in the Lorentz coordinate system of $\mathbf{e}$. Recall that these quantities are not Lorentz invariant. Thus two different inertial observers will, in general, observe two different Electric and Magnetic fields. However all inertial observers will observe that the Maxwell's equations hold with respect to their observed fields. Furthermore from (3.4) we see that, in every Loretntz coordinate system the classical force acting on the charged particle of charge q will be given by the *Lorentz force*

$$\mathbf{f}_q = q\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right). \tag{3.15}$$

We also see that conservation of charge expressed by equation (3.9) takes the usual form[1]

$$\frac{\partial \rho}{\partial t} + \nabla \mathbf{J} = 0. \tag{3.16}$$

We see that equation 3.10 takes the form

$$\frac{d}{dt}\int_{V^2(t)} \mathbf{B}\cdot d\mathbf{s} = -\int_{\partial V^2(t)} c\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right). \tag{3.17}$$

Let us now see how the electromagnetic form $\mathscr{F}$ transforms under Lorentz transformations. That is how Electric and Magnetic fields, $\mathbf{E}$ and $\mathbf{B}$, in one inertial frame relate to the electric and magnetic fields observed in another inertial frame. Consider a Lorentz transformation of a boost in the x-direction.

$$\begin{aligned} t &= \gamma\left(\tilde{t} + \frac{v}{c^2}\tilde{x}^1\right), \\ x^1 &= \gamma(\tilde{x}^1 + v\tilde{t}), \\ x^2 &= \tilde{x}^2, \end{aligned}$$

[1] The continuity equation.

$$x^3 = \tilde{x}^3.$$

Writing $\mathscr{F}$ in the inertial frame $(\tilde{t}, \tilde{x}^1, \tilde{x}^2, \tilde{x}^3)$ and re-writing in the form

$$\mathscr{F} = \tilde{\mathscr{E}} \wedge d\tilde{t} + \frac{1}{c}\tilde{\mathscr{B}}$$

we have

$$\begin{aligned}
\tilde{E}_1 &= E_1, \\
\tilde{E}_2 &= \gamma\left(E_2 - \frac{v}{c}B_3\right), \\
\tilde{E}_3 &= \gamma\left(E_3 + \frac{v}{c}B_2\right).
\end{aligned}$$

$$\begin{aligned}
\tilde{B}_1 &= B_1, \\
\tilde{B}_2 &= \gamma\left(B_2 + \frac{v}{c}E_3\right), \\
\tilde{B}_3 &= \gamma\left(B_3 - \frac{v}{c}E_2\right).
\end{aligned}$$

Thus we see that though an observer in (t, x^1, x^2, x^3) would observe that $B \equiv 0$ an observer in an inertial frame $(t, \tilde{x}^1, \tilde{x}^2, \tilde{x}^3)$ would observe that $\tilde{B} \neq 0$. Thus demonstrating that electric and magnetic fields are not two different things but different manifestations of the same thing. Furthermore it can be shown that $\text{vol}^4 = (1/c^3)\, dt \wedge dx^1 \wedge dx^2 \wedge dx^3 = (1/c^3)\, d\tilde{t} \wedge d\tilde{x}^1 \wedge d\tilde{x}^2 \wedge d\tilde{x}^3$.

In summary every inertial observer will experience an Electric field $\mathbf{E}$ and a Magnetic field $\mathbf{B}$ and observe that they satisfy the Maxwell's equations. They will also observe that the force acting on a charged particle of charge q is given by the Lorentz force (7.48). However, in general, different inertial observers will not agree on the value of the Electric field $\mathbf{E}$, Magnetic field $\mathbf{B}$, and the Lorentz force $\mathbf{f}_q$.

Furthermore as a consequence of the Maxwell's laws, we see from (7.50) and (7.53) that all inertial observers will see that $(||\mathbf{B}||^2 - ||\mathbf{E}||^2)$ and $\mathbf{E} \cdot \mathbf{B}$ will have the same value and that equations (3.9) and (3.10) will hold.

3.1 The potential form of the Maxwell's equations

We point out that (3.5)–(3.8) can be combined to give

$$\begin{aligned}
\frac{1}{c^2}\frac{\partial^2 \mathscr{E}_s}{\partial t^2} &= -\star \mathbf{d}_s \star \mathbf{d}_s \mathscr{E}_s - \frac{4\pi}{c^2}\frac{\partial \star \mathscr{J}_s}{\partial t}, \\
\frac{1}{c^2}\frac{\partial^2 \mathscr{B}_s}{\partial t^2} &= -\mathbf{d}_s \star \mathbf{d}_s \star \mathscr{B}_s + \frac{4\pi}{c}\mathbf{d}_s \mathscr{J}_s.
\end{aligned}$$

Using the *Laplace operator* $\Delta = (\star\mathbf{d}_s\star\mathbf{d}_s - \mathbf{d}_s\star\mathbf{d}_s\star)$ the above becomes

$$\begin{aligned}\frac{1}{c^2}\frac{\partial^2\mathscr{E}_s}{\partial t^2} &= -\Delta\mathscr{E}_s - \mathbf{d}_s\star\mathbf{d}_s\star\mathscr{E}_s - \frac{4\pi}{c^2}\frac{\partial\star\mathscr{J}_s}{\partial t},\\ \frac{1}{c^2}\frac{\partial^2\mathscr{B}_s}{\partial t^2} &= \Delta\mathscr{B}_s - \star\mathbf{d}_s\star\mathbf{d}_s\mathscr{B}_s + \frac{4\pi}{c}\mathbf{d}_s\mathscr{J}_s.\end{aligned}$$

Hence finally from the two Maxwell's equations (3.7)–(3.8) we have

$$\frac{1}{c^2}\frac{\partial^2\mathscr{E}_s}{\partial t^2} = -\Delta\mathscr{E}_s - 4\pi\mathbf{d}_s\star\mathscr{D}_s - \frac{4\pi}{c^2}\frac{\partial\star\mathscr{J}_s}{\partial t}, \tag{3.18}$$

$$\frac{1}{c^2}\frac{\partial^2\mathscr{B}_s}{\partial t^2} = \Delta\mathscr{B}_s + \frac{4\pi}{c}\mathbf{d}_s\mathscr{J}_s. \tag{3.19}$$

In usual vector notation these take the form

$$\begin{aligned}\frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} &= -\nabla\times\nabla\times\mathbf{E} - \frac{4\pi}{c}\frac{\partial\mathbf{J}}{\partial t} = \nabla^2\mathbf{E} - \nabla(\nabla\cdot\mathbf{E}) - \frac{4\pi}{c}\frac{\partial\mathbf{J}}{\partial t},\\ \frac{1}{c^2}\frac{\partial^2\mathbf{B}}{\partial t^2} &= -\nabla\times\nabla\times\mathbf{B} + \frac{4\pi}{c}\nabla\times\mathbf{J} = \nabla^2\mathbf{B} - \nabla(\nabla\cdot\mathbf{B}) + \frac{4\pi}{c}\nabla\times\mathbf{J},\end{aligned}$$

where we have used the identity

$$\nabla\times(\nabla\times\mathbf{V}) = \nabla(\nabla\cdot\mathbf{V}) - \nabla^2\mathbf{V}.$$

Thus the two Maxwell's equations (3.13)–(3.14) give that

$$\frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} = \nabla^2\mathbf{E} - 4\pi\nabla\rho - \frac{4\pi}{c}\frac{\partial\mathbf{J}}{\partial t}, \tag{3.20}$$

$$\frac{1}{c^2}\frac{\partial^2\mathbf{B}}{\partial t^2} = \nabla^2\mathbf{B} + \frac{4\pi}{c}\nabla\times\mathbf{J}, \tag{3.21}$$

Since the 2-form $\mathscr{F}$ is closed, $d\mathscr{F} = 0$. Then the Poincare lemma implies that $\mathscr{F}$ is locally exact. That is there exists a 1-form

$$\mathscr{A} = \phi\, dt + \frac{1}{c}\mathscr{A}_s, \tag{3.22}$$

such that $\mathscr{F} = d\mathscr{A}$. The function ϕ is called the scalar potential the spatial 1-form $\mathscr{A}_s = A_1 dx^1 + A_2 dx^2 + A_3 dx^3$ is called the vector potential of the field.

Since

$$d\mathscr{A} = \left(\mathbf{d}_s\phi - \frac{1}{c}\frac{\partial\mathscr{A}_s}{\partial t}\right)\wedge dt + \frac{1}{c}\mathbf{d}_s\mathscr{A}_s.$$

we have that

$$\left(\mathbf{d}_s\phi - \frac{1}{c}\frac{\partial\mathscr{A}_s}{\partial t}\right) = \mathscr{E}_s, \tag{3.23}$$

$$\mathbf{d}_s\mathscr{A}_s = \mathscr{B}_s. \tag{3.24}$$

Thus from (7.55)–(7.57) we have that the Maxwell's equations $d\mathscr{F}=0$ and $d*\mathscr{F}=4\pi\mathscr{J}$ reduce to

$$\frac{1}{c}\mathbf{d}_s\star\frac{\partial\mathscr{A}_s}{\partial t}-\mathbf{d}_s\star\mathbf{d}_s\phi=-4\pi\mathscr{D}_s \tag{3.25}$$

$$\frac{1}{c^2}\star\frac{\partial^2\mathscr{A}_s}{\partial t^2}-\left(\frac{1}{c}\star\frac{\partial\mathbf{d}_s\phi}{\partial t}-\mathbf{d}_s\star\mathbf{d}_s\mathscr{A}_s\right)=\frac{4\pi}{c}\mathscr{J}_s, \tag{3.26}$$

The potential $\mathscr{A}$ such that $d\mathscr{A}=\mathscr{F}$ is not unique since any $\mathscr{A}+df$ also satisfies $d\mathscr{A}=\mathscr{F}$. Thus to ensure uniqueness it is additionally required that $\mathscr{A}$ be such that $d*\mathscr{A}=0$. Since

$$*\mathscr{A}=\frac{1}{c^2}\phi\,\mathrm{vol}_s+\frac{1}{c}\star\mathscr{A}_s\wedge dt,$$

$$d*\mathscr{A}=\frac{1}{c}\left(-\frac{1}{c}\frac{\partial\phi}{dt}\mathrm{vol}_s+\mathbf{d}_s\star\mathscr{A}_s\right)\wedge dt,$$

where we have set $\mathrm{vol}_s=dx^1\wedge dx^2\wedge dx^3$, we see that the requirement $d*\mathscr{A}=0$ gives $\frac{1}{c}\frac{\partial\phi}{dt}\mathrm{vol}_s-\mathbf{d}_s\star\mathscr{A}_s=0$ and taking the $\star$ we have

$$\frac{1}{c}\frac{\partial\phi}{dt}-\star\mathbf{d}_s\star\mathscr{A}_s=0. \tag{3.27}$$

This is usually referred to as the *Lorentz Gauge* condition and says that the potentials are divergence free. The vector version of this expression is

$$\frac{1}{c}\frac{\partial\phi}{dt}-\nabla\cdot\mathbf{A}=0. \tag{3.28}$$

Subjected to the Lorentz gauge (3.25) and (3.26) become

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial^2 t}-\Delta\phi=-4\pi\star\mathscr{D}_s \tag{3.29}$$

$$\frac{1}{c^2}\frac{\partial^2\mathscr{A}_s}{\partial t^2}+\Delta\mathscr{A}_s=\frac{4\pi}{c}\star\mathscr{J}_s. \tag{3.30}$$

where the operator $\Delta=(\star\mathbf{d}_s\star\mathbf{d}_s-\mathbf{d}_s\star\mathbf{d}_s\star)$ is called the *Laplace operator*. The vector version of the Laplace operator is $\Delta\mathbf{A}\triangleq\nabla^2\mathbf{A}\triangleq\nabla\cdot(\nabla\mathbf{A})-\nabla\times\nabla\times\mathbf{A}$ and $\Delta\phi\triangleq\nabla^2\phi\triangleq\nabla\cdot(\nabla\phi)$. Thus, the vector version of the gauge condition (3.28), implies that the vector version of the above equations are

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial^2 t}-\nabla^2\phi=-4\pi\rho^e \tag{3.31}$$

$$\frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2}-\nabla^2\mathbf{A}=\frac{4\pi}{c}\mathbf{J}. \tag{3.32}$$

3.2 The vacuum solution to Maxwell's equations

In a vacuum $\mathbf{J} = 0, \rho = 0$ and we see that in Lorentz coordinates the plane wave

$$\mathbf{E} = \mathbf{E_k} e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})},$$
$$\mathbf{B} = \mathbf{B_k} e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}$$

where $\mathbf{E_k}, \mathbf{B_k} \in \mathbb{C}^3$ are complex constants, satisfy the wave equations (3.20)–(3.21) if and only if $\mathbf{k}$ satisfies

$$\omega^2 - c^2 k^2 = 0,$$

where $k^2 = ||\mathbf{k}||^2$. Thus the solutions correspond to plane waves with group velocity equal to c. By substituting in the Maxwell's equations we can show that the cyclic permutations of the identity

$$\mathbf{E_k} \times \mathbf{B_k} = \frac{\mathbf{k}}{k}, \tag{3.33}$$

hold.

We will consider a wave for some fixed $\mathbf{k}$. Let us pick the coordinate system $\mathbf{e}$ so that the wave is propagating in the $\mathbf{e}_3$ direction. Let $r = (x, y, z)$ in this coordinate system. Then a wave propagating in the $\mathbf{e}_3$ direction corresponds to $\mathbf{k} = (0, 0, k)$ and has the form

$$\mathbf{E} = \mathbf{E_k} e^{i(\omega t - kz)},$$
$$\mathbf{B} = \mathbf{B_k} e^{i(\omega t - kz)}.$$

The wave orthogonality condition (3.33) then implies that $\mathbf{E_k} = (E_{kx} e^{i\phi_x}, E_{ky} e^{i\phi_y}, 0)$ where $E_{kx}, E_{ky} \in \mathbb{R}$ are real constants. Since $\mathbf{B_k}$ is completely determined by $\mathbf{E_k}$ and $\mathbf{k}$ we will only consider $\mathbf{E_k}$. Thus we have

$$\mathbf{E} = (E_{kx} e^{i(\omega t - kz + \phi_x)}, E_{ky} e^{i(\omega t - kz + \phi_y)}, 0),$$

and hence that

$$\Re(\mathbf{E}) = (E_{kx} \cos(\omega t - kz + \phi_x), E_{ky} \cos(\omega t - kz + \phi_y), 0).$$

When $\phi_x = \phi_y$ we say that the light is *linearly polarized* in the $(E_{kx}, E_{ky}, 0)$ direction and when $\phi_x - \phi_y = \pm\pi/2$ we say that the light is *eliptically polarized*. When the plus is chosen the light is rotating in the left hand direction while when the minus is chosen the light is rotating in the right hand direction. In the special case where $E_{kx} = E_{ky} = E_0$ the polarization is said to be circular. Thus a circularly polarized light is given by

$$\Re(\mathbf{E}_{c\pm}) = E_0(\cos(\omega t - kz), \mp\sin(\omega t - kz), 0).$$

Note that

$$\frac{1}{2}\left(\Re(\mathbf{E}_{c+}) + \Re(\mathbf{E}_{c-})\right) = E_0(1, 0, 0)\cos(\omega t - kz).$$

and hence that the sum of a left oriented circularly polarized light plus a right oriented circularly polarized light results in a linearly polarized light.

The angular momentum of circularly polarized light is defined to be

$$\mathbf{J} = \pm e\,||\mathbf{E_k}||^2 \frac{\mathbf{k}}{||\mathbf{k}||}.$$

3.3 Physical Significance of the Electromagnetic Potential

Consider the motion of a classical particle along a curve $c : I \to \mathbb{R}^4$ in the absence and presence of an electric field $\mathscr{F} = d(\phi dt + \mathscr{A}_s)$. The Feynman path integrand in the absence of the field is given by

$$\Phi(c(s)) = \exp\left\{ \int_{c(s)} \left(\frac{i}{\hbar}\right) \Lambda \right\},$$

while in the presence of the field is given by

$$\Phi^e(c(s)) = \exp\left\{ \int_{c(s)} \left(\frac{i}{\hbar}\right) \Lambda^e \right\} = \exp\left\{ \int_{c(s)} \left(\frac{i}{\hbar}\right) (\Lambda + e\mathscr{A}) \right\}.$$

Then for a path $c : I \to \mathbb{R}^4$ that is given by $(t, \gamma(s))$ where γ is the boundary of some compact surface S

$$\begin{aligned}
\Phi^e(c(s)) &= \exp\left\{ \int_{c(s)} \left(\frac{i}{\hbar}\right) \Lambda^e \right\} = \exp\left\{ \int_{c(s)} \left(\frac{i}{\hbar}\right) (\Lambda + e\mathscr{A}) \right\} \\
&= \Phi(c(s)) \exp\left\{ \int_{c(s)} \left(\frac{i}{\hbar}\right) (\phi dt + e\mathscr{A}_s) \right\}, \\
&= \Phi(c(s)) \exp\left\{ \int_{\gamma(s)} \left(\frac{ie}{\hbar}\right) (\mathscr{A}_s) \right\} \\
&= \Phi(c(s)) \exp\left\{ \int_{S} \left(\frac{ie}{\hbar}\right) \mathscr{B}_s \right\}.
\end{aligned}$$

For S that contains a cross section of a tightly wound infinitely long current carrying solenoid $\int_S \mathscr{B}_s = b$ and thus for such a case we have

$$\Phi^e(c(s)) = \Phi(c(s)) \exp\left\{ \int_{S} \left(\frac{ie}{\hbar}\right) \mathscr{B}_s \right\} = \Phi(c(s)) \exp\left\{ \left(\frac{ieb}{\hbar}\right) \right\}.$$

Hence integrating over all cyclic paths we have

$$\int \Phi^e(c(s)) = \exp\left\{ \left(\frac{ieb}{\hbar}\right) \right\} \int \Phi(c(s)).$$

Chapter 4
Variational Calculus and Hamilton's Equations

Consider a curve $c : I \to M^n$ where $I = [s_1, s_2]$ and a 1-form Λ. We define the *action* of Λ over $c(s)$ as

$$S(c(s)) = \int_{c(s)} \Lambda = \int_I c^*\Lambda = \int_I \Lambda\left(\frac{dc}{ds}\right) ds$$

Consider a variation of the curve $c : I \to M^n$ given by $c(\tau, \cdot) : I \times [-1, 1] \to M^n$ such that $c(\tau, 0) = c(s)$. Then denote by $J(\tau, \alpha), T(\tau, \alpha) \in T_{c(\tau,\alpha)}M^n$ by

$$J(\tau, \alpha) = c_*\left(\frac{\partial}{\partial \alpha}\right),$$
$$T(\tau, \alpha) = c_*\left(\frac{\partial}{\partial \tau}\right)$$

For a fixed α denote the curve $c_\alpha(\tau) = c(\tau, \alpha)$. Consider the action of Λ along the curve $c_\alpha(\tau) = c(\tau, \alpha)$ given by

$$S(\alpha) = \int_{c_\alpha(\tau)} \Lambda = \int_I c_\alpha^*\Lambda = \int_I \Lambda(T(\tau, \alpha))d\tau.$$

$$\begin{aligned}
\frac{dS}{d\alpha} &= \int_{c_\alpha(\tau)} \left(\frac{\partial \Lambda}{\partial \alpha} + \mathscr{L}_{J(\tau,\alpha)}\Lambda\right) = \int_{c_\alpha(\tau)} \left(\frac{\partial \Lambda}{\partial \alpha} + di_{J(\tau,\alpha)}\Lambda + i_{J(\tau,\alpha)}d\Lambda\right) \\
&= \int_{\partial c_\alpha(\tau)} i_{J(\tau,\alpha)}\Lambda + \int_{c_\alpha(\tau)} \left(\frac{\partial \Lambda}{\partial \alpha} + i_{J(\tau,\alpha)}d\Lambda\right) \\
&= (\Lambda(J(\tau_2, \alpha)) - \Lambda(J(\tau_1, \alpha))) + \int_{c_\alpha(\tau)} \left(\frac{\partial \Lambda}{\partial \alpha} + i_{J(\tau,\alpha)}d\Lambda\right) \qquad (4.1)
\end{aligned}$$

Since Λ does not depend on the variation parameter α and we only consider variations that vanish at the end points. That is $J(\tau_1, 0)) = J(\tau_2, 0) = 0$. Then we have

$$\left.\frac{dS}{d\alpha}\right|_{\alpha=0} = \int_{c(s)} i_{J(\tau,0)}d\Lambda \qquad (4.2)$$

If $S(c(t))$ is stationary at $c(t)$ then $\left.\frac{dS}{d\alpha}\right|_{\alpha=0} = 0$ for all variations $J(\tau, 0)$ of $c(t)$ that vanish at the end points. Which implies that $i_J d\Lambda = 0$ for all J that are tangent to a variational surface that contains $c(t,0)=c(s)$. Thus, since the tangent to the curve should also be on such a surface we should necessarily have $i_{\dot{c}}d\Lambda = 0$. In summary we have proved the following theorem:

Theorem 4.1. *Consider a curve $c : I \to M^n$ where $I=[s_1,s_2]$ and a 1-form Λ. The statement that*

$$S(c(s)) \triangleq \int_{c(s)} \Lambda$$

is stationary at $c(c)$ for all variations of $c(s)$ that vanish at the end points is equivalent to the statement that $d\Lambda \triangleq \Omega \equiv 0$ when restricted to $c(s)$. That is $i_{q'}\Omega = 0$ where $q'(s) = \dfrac{dc}{ds}$.

4.1 Hamilton's Equations

Consider $\mathbb{R}\times T^*M$ and the 1-form

$$\Lambda = p_i dq^i - H dt. \tag{4.3}$$

Then for the motion of a classical particle describe by the curve $c : I \to \mathbb{R}\times T^*M$ that takes $t \to c(t) = (t, q(t), p(t)) \in \mathbb{R}\times T^*M$ we have that the action along the motion is given by

$$S(c(t)) \triangleq \int_{c(t)} \Lambda = \int_I i_{\dot{c}}\Lambda \, dt = \int_I L(q,\dot{q})\, dt$$

where $i_{\dot{c}}\Lambda = p\cdot\dot{q} - H = \frac{1}{2}m||\dot{q}||^2 - V(q) = L(q,\dot{q})$ is defined to be the *Lagrangian* of the motion.

We see that

$$\begin{aligned}\Omega \triangleq d\Lambda &= dp_i \wedge dq^i + dt\wedge dH &(4.4)\\ &= \frac{\partial H}{\partial q^i} dt\wedge dq^i + \left(\frac{\partial H}{\partial p_i}dt - dq^i\right)\wedge dp_i. &(4.5)\end{aligned}$$

Then

$$\begin{aligned}i_{\dot{c}}\Omega &= -\left(\frac{\partial H}{\partial q^i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i\right)dt + \left(\frac{\partial H}{\partial q^i} + \dot{p}_i\right)dq^i + \left(\frac{\partial H}{\partial p_i} - \dot{q}^i\right)dp_i\\ &= \left(\frac{\partial H}{\partial t} - \frac{dH}{dt}\right)dt + \left(\frac{\partial H}{\partial q^i} + \dot{p}_i\right)dq^i + \left(\frac{\partial H}{\partial p_i} - \dot{q}^i\right)dp_i.\end{aligned}$$

This shows that $i_{\dot{c}}\Omega = 0$ is equivalent to the set of equations

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}, \tag{4.6}$$

$$\dot{q} = \frac{\partial H}{\partial p}, \tag{4.7}$$

$$\dot{p} = -\frac{\partial H}{\partial q}. \tag{4.8}$$

These are the *Hmilton's equations* for a system of interacting particles with Hamiltonian $H : \mathbb{R} \times T^*M \to \mathbb{R}$. Thus we have the following theorem.

Theorem 4.2. *Let $H : \mathbb{R} \times T^*M \to \mathbb{R}$ to be a smooth function called the Hamiltonian, $c : [t_1, t_2] \to \mathbb{R} \times T^*M$ to be a smooth curve that takes the coordinate form $c(t) = (t, q(t), p(t))$, Λ to be the energy-momentum 1-form defined by*

$$\Lambda = p_i dq^i - H dt, \tag{4.9}$$

and $S(c(t))$ to bethe action of the system along $c(t)$

$$S(c(t)) = \int_{c(t)} \Lambda = \int_{t_1}^{t_2} L(q, \dot{q}) dt. \tag{4.10}$$

The curve $c(t)$ is a stationery point of the action if and only if $c(t)$ satisfies the Hamilton's equations (7.68) – (7.70).

The vector field

$$X_H = \frac{\partial}{\partial t} + \frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p},$$

is defined to be the Hamiltonian vector field on $\mathbb{R} \times T^*M$ that is associated with H. Since $i_{X_H}\Omega = 0$ and $d\Omega = 0$ it also follows that

$$\mathscr{L}_{X_H}\Omega = 0. \tag{4.11}$$

4.2 Hamilton's Equations in the Presence of an Electromagnetic Field

Consider a particle with an associated Hamiltonian H in the absence of an electro-magnetic field. If in addition the particle has a charge e then it is also influenced by the electromagnetic field $\mathscr{F} = \mathscr{E}_s \wedge dt + \frac{1}{c}\mathscr{B}_s$ where $\mathscr{E}_s$ is the electrostatic 1-form and $\mathscr{B}_s$ is the magnetic 2-form associated with the classical fields $\mathbf{E} = (E_1, E_2, E_3)$ and $\mathbf{B} = (B_1, B_2, B_3)$. Here both $\mathscr{E}_s$ and $\mathscr{B}_s$ do not contain dt terms.

Since from Maxwell's equations we have $d\mathscr{F} = 0$ we are motivated to define a new action by augmenting Λ with $\mathscr{A}$ as follows.

$$\Lambda^e \triangleq \Lambda + e\mathscr{A} = (p_i + eA_i)dq^i - (H - e\phi)dt = p_i^e dq^i - H^e dt. \tag{4.12}$$

Then

$$\Omega^e \triangleq d\Lambda^e = \Omega + e\mathscr{F}. \tag{4.13}$$

Computing $i_{\dot{c}}\Omega^e$ we find

$$\begin{aligned} i_{\dot{c}}\Omega^e &= i_{\dot{c}}\Omega + e i_{\dot{c}}\mathscr{F} \\ &= i_{\dot{c}}\Omega + e(i_{\dot{q}}\mathscr{E}_s)dt - e\left(\mathscr{E}_s - \frac{1}{c} i_{\dot{q}}\mathscr{B}_s\right). \end{aligned}$$

Writing this out in Euclidean coordinates for space we have

$$i_{X_H}\Omega^e = \left(\frac{\partial H}{\partial t} - \frac{dH}{dt} + e\,(\dot{q}\cdot\mathbf{E})\right)dt + \left(\frac{\partial H}{\partial q} + \dot{p}\right)dq + \left(\frac{\partial H}{\partial p} - \dot{q}\right)dp - e\,(\mathbf{E} + (\dot{q}\times\mathbf{B}))\,dq.$$

If the action $S^e(c(t)) = \int_{c(t)} \Lambda^e$ is stationary along the curve $c(t) = (t, q(t), p(t))$ then we have seen that $c(t)$ should satisfy the condition $i_X\Omega^e = 0$ and hence

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + e\,(\dot{q}\cdot\mathbf{E}), \tag{4.14}$$

$$\dot{q} = \frac{\partial H}{\partial p}, \tag{4.15}$$

$$\dot{p} = -\frac{\partial H}{\partial q} + e\,(\mathbf{E} + (\dot{q}\times\mathbf{B})). \tag{4.16}$$

In fact experiments have verified that the behaviour of a classical particle influenced by the presence of an electromagnetic field is sufficiently accurately defined by these equations.

Since $d\mathscr{F} = 0$ we see that there exists a potential $\mathscr{A} = \phi dt + \frac{1}{c}\mathscr{A}_s$ such that $d\mathscr{A} = \mathscr{F}$. Note that this potential is unique only up to an addition by a differential of a function. Then we have

$$\begin{aligned} i_{\dot{c}}\Omega^e &= i_{\dot{c}}\Omega + e i_{\dot{c}}\mathscr{F} \\ &= i_{\dot{c}}\Omega + e(i_{\dot{q}}\mathscr{E}_s)dt - e(\mathscr{E}_s - i_{\dot{q}}\mathscr{B}_s) \\ &= i_{\dot{c}}\Omega + e\left(i_{\dot{q}}\mathbf{d}_s\phi - \frac{\partial i_{\dot{q}}\mathscr{A}_s}{\partial t}\right)dt - e\left(\left(\mathbf{d}_s\phi - \frac{\partial \mathscr{A}_s}{\partial t}\right) - i_{\dot{q}}\mathbf{d}_s\mathscr{A}_s\right) \\ &= i_{\dot{c}}\Omega + e\left(\frac{d\phi}{dt} - \frac{\partial \phi}{\partial t} - \frac{\partial i_{\dot{q}}\mathscr{A}_s}{\partial t}\right)dt - e\left(\left(\mathbf{d}_s\phi - \frac{\partial \mathscr{A}_s}{\partial t}\right) - i_{\dot{q}}\mathbf{d}_s\mathscr{A}_s\right). \end{aligned}$$

Define where $H^e \triangleq H(t, q, p^e - eA) - e\phi(t, q)$ and using the fact that $\dot{q} = \dfrac{\partial H}{\partial p}$ we see that

$$\begin{aligned} \frac{\partial H^e}{\partial t} &= \frac{\partial}{\partial t}(H - e\phi) - e\frac{\partial H}{\partial p}\frac{\partial A}{\partial t} = \frac{\partial}{\partial t}(H - e\phi) - e\frac{\partial}{\partial t}(A\cdot\dot{q}), \\ \frac{\partial H^e}{\partial q} &= \frac{\partial}{\partial q}(H - e\phi) - e\frac{\partial A}{\partial q}\frac{\partial H}{\partial p} = \frac{\partial}{\partial q}(H - e\phi) - e\frac{\partial}{\partial q}(A\cdot\dot{q}). \end{aligned}$$

Thus by direct computation we find that

$$\begin{aligned}
i_{\dot{c}}\Omega^e &= \left(\frac{\partial}{\partial t}(H - e\phi - e\mathbf{A}\cdot\dot{q}) - \frac{d}{dt}(H - e\phi)\right)dt + \left(-e\left(\frac{\partial \phi}{\partial q} - \frac{\partial \mathbf{A}}{\partial t} + \dot{q}\times\nabla\mathbf{A}\right) + \frac{\partial H}{\partial q} + \dot{p}\right)dq + \left(\frac{\partial H}{\partial p} - \dot{q}\right)dp \\
&= \left(\frac{\partial}{\partial t}(H - e\phi - e\mathbf{A}\cdot\dot{q}) - \frac{d}{dt}(H - e\phi)\right)dt + \left(-e\left(-\frac{\partial \mathbf{A}}{\partial t} + \frac{d\mathbf{A}}{dt} + \dot{q}\times\nabla\mathbf{A}\right) + \frac{\partial}{\partial q}(H - e\phi) + \frac{d}{dt}(p + e\mathbf{A})\right)dq + \left(\frac{\partial H}{\partial p} - \dot{q}\right)dp \\
&= \left(\frac{\partial}{\partial t}(H - e\phi - e\mathbf{A}\cdot\dot{q}) - \frac{d}{dt}(H - e\phi)\right)dt + \left(-e\left(-\frac{\partial \mathbf{A}}{\partial t} + \frac{d\mathbf{A}}{dt} + \dot{q}\times\nabla\mathbf{A} - \frac{\partial}{\partial q}(A\cdot\dot{q})\right) + \frac{\partial}{\partial q}(H - e(A\cdot\dot{q}) - e\phi) + \frac{d}{dt}(p + e\mathbf{A})\right)dq + \left(\frac{\partial H}{\partial p} - \dot{q}\right)dp \\
&= \left(\frac{\partial H^e}{\partial t} - \frac{dH^e}{dt}\right)dt + \left(\frac{\partial H^e}{\partial q} + \frac{dp^e}{dt}\right)dq + \left(\frac{\partial H^e}{\partial p^e} - \dot{q}\right)dp.
\end{aligned}$$

where we have used the fact

$$0 = \left(-\frac{\partial \mathbf{A}}{\partial t} + \frac{d\mathbf{A}}{dt} + \dot{q}\times\nabla\mathbf{A} - \frac{\partial}{\partial q}(\mathbf{A}\cdot\dot{q})\right).$$

Thus we see that equations the motion of a classical particle under the influence of an electromagnetic field is given by the Hamilton's equations

$$\frac{dH^e}{dt} = \frac{\partial H^e}{\partial t}, \tag{4.17}$$

$$\dot{q} = \frac{\partial H^e}{\partial p^e}, \tag{4.18}$$

$$\dot{p}^e = -\frac{\partial H^e}{\partial q}, \tag{4.19}$$

with electromagnetic Hamiltonian $H^e \triangleq H(q, p^e - e\mathbf{A}) - e\phi$ and modified momentum $p^e = p + e\mathbf{A}$.

4.3 Poisson Manifolds and Hamiltonian Dynamcs

Let $\mathscr{F}(M)$ denote the space of infinitely differentiable functions on a smooth n-dimensional manifold M. A bilinear operators $\{\cdot,\cdot\} : \mathscr{F}\times\mathscr{F} \to \mathbb{R}$ such that $\mathscr{F}$ is a Lie algebra with Lie bracket $\{\cdot,\cdot\}$ and $\{\cdot,\cdot\}$ is a derivation on $\mathscr{F}$. That is

$$\{fg, h\} = f\{g,h\} + g\{f,h\} \quad \forall \quad f,g \in \mathscr{F}$$

Let $\mathscr{X}(M)$ denote the space of infinitely differentiable vector fields on M. A vector field $X_H \in \mathscr{X}$ is said to be Hamiltonian with respect to $H \in \mathscr{F}$ if

$$X_H(f) = \{f, H\} \quad \forall \quad f \in \mathscr{F}.$$

Let G be a Lie group and $\phi : G\times M \to M$ be a left action of G on M. The action is said to be canonical if

$$\phi_g^*\{f,g\} = \{\phi_g^* f, \phi_g^* g\} \quad \forall \quad f,g \in \mathscr{F}.$$

Define for $\zeta \in \mathscr{G}$ the vector field $\zeta_M \in \mathscr{X}$ that is give by

$$\zeta_M(m) \triangleq \frac{d}{dt}\Big|_{t=0}\phi_{\exp \zeta t}(m) \quad \forall \quad m \in M \; .$$

Define the map $J : \mathscr{G} \to \mathscr{F}$ such that for each $\zeta \in \mathscr{G}$ the vector field ζ_M is Hamiltonian with respect to J_ζ. That is

$$X_{J_\zeta} = \zeta_M.$$

or alternatively

$$\zeta_M(f) = \{f, J_\zeta\} \quad \forall \quad f \in \mathscr{F}.$$

Note that J_ζ is unique only upto an additive Casimir function and is thus a linear function. The map $\mathbf{J} : M \to \mathscr{G}^*$ defined such that

$$\langle \mathbf{J}(m), \zeta \rangle \triangleq J_\zeta(m),$$

is called a *momentum map* of the action ϕ.

If $\zeta_M(H) = 0$ for all $\zeta \in \mathscr{G}$ then $\{H, J_\zeta\} = 0$ for all $\zeta \in \mathscr{G}$. Hence $X_H(J_\zeta)$ for all $\zeta \in \mathscr{G}$. Thus $J_\zeta \in \mathscr{F}$ is conserved along the flow of X_H for all $J \in \mathscr{G}$ and hence $\mathbf{J}$ is conserved along the flow of X_H. This is Noether's theorem in the Hamiltonian setting.

Theorem 4.3. Noether's Theorem: *If $\phi : G \times M \to M$ is a canonical left action of G on M and $H \in \mathscr{F}$ is invariant under the action of ϕ then the momentum map $\mathbf{J}$ corresponding to the action of ϕ is conserved under the flow of X_H.*

Chapter 5
Schrödinger's equation in an Electromagnetic Field

Consider the smooth manifold M and the k-dimensional vector space $\mathbb{K}$. Let $\{U,V,\cdots\}$ be an open cover of M. The 4-tuple $\{E,M,\mathbb{K},\pi\}$ is called a rank k vector bundle with fibre equal to $\mathbb{K}$; if for each $U,V \subset M$ such that $U \cap V \neq \{\emptyset\}$ there exists diffeomorphisms $\phi_U : \pi^{-1}(U) \to U \times \mathbb{K}$, $\phi_V : \pi^{-1}(V) \to V \times \mathbb{K}$ that satisfy the property that for each $p \in \pi^{-1}(U \cap V)$ the fiber coordinates k^U and k^V, given by $\phi_U(p) = (m^V, k^U)$ and $\phi_V(p) = (m^V, k^V)$ respectively, are related by $k^V = c_{VU}k^U$ for some $c_{VU} : U \cap V \to GL(\mathbb{K})$ that satisfy the properties $c_{UV}^{-1} = c_{VU}$ and $c_{VU}c_{UW}c_{WV} = id$. The $c_{VU} \in GL(\mathbb{K})$ are called the transition maps.

Then a section of the bundle $\psi : M \to E$ is defined by a collection of complex valued locally defined $\mathbb{K}$ valued functions $\{\psi^U, \psi^V, \cdots\}$ in each patch $\{U, \phi_U\}$ such that they satisfy the property $\psi^V = c_{VU}\psi^U$ in overlaps.

Let $\Gamma(E)$ denote the space of all possible smooth sections of the bundle. Then a connection ∇ of the bundle is a bilinear map $\nabla : TM \times \Gamma(E) \to \Gamma(E)$ that satisfies the Leibnitz rule. That is it maps sections to section valued 1-forms in a linear fashion such that $\nabla \psi f = (\nabla \psi) f + \psi \otimes df$ holds. In each coordinate patch we then have

$$\nabla^U \psi^U \triangleq \omega^U \psi^U + d\psi^U, \tag{5.1}$$

where the collection of $GL(\mathbb{K})$ valued 1-forms $\{\omega^U, \omega^V, \cdots\}$ satisfy the property

$$\omega^V = c_{UV}^{-1}\omega^U c_{UV} + c_{UV}^{-1} dc_{UV}, \tag{5.2}$$

that ensures $\nabla^V \psi^V = c_{VU}\nabla^U \psi^U$ in overlaps. Note that from (5.1) we see that

$$\nabla_{\partial_i} = \partial_i + \omega(\partial_i). \tag{5.3}$$

The curvature of the connection is given by the collection of 2-forms $\{\theta^U, \theta^V, \cdots\}$, which in a coordinate patch are given by

$$\theta^U \triangleq \omega^U \wedge \omega^U + d\omega^U. \tag{5.4}$$

The local coordinate versions are related to each other by

$$\theta^V = c_{UV}^{-1}\theta^U c_{UV}. \tag{5.5}$$

In the special case of a complex line bundle $\{E,M,\mathbb{C},\pi\}$ we see that $GL(\mathbb{C})=\mathbb{C}$ and hence that $\theta^U=d\omega^U$ and

$$\omega^V = c_{UV}^{-1}\omega^U c_{UV} + c_{UV}^{-1}dc_{UV} = \omega^U + d\ln c_{UV}, \tag{5.6}$$
$$\theta^V = c_{UV}^{-1}\theta^U c_{UV} = \theta^U. \tag{5.7}$$

In particular we have that the curvature is a true two form on M. Thus any exact two form θ on M defines a connection on $\{E,M,\mathbb{C},\pi\}$ as follows. Since θ is exact there exists a 1-form ω^U such that $\theta = d\omega^U$. Note that ω^U is not unique since $\omega^V \triangleq \omega^U + df_{UV}$ for any complex function f_{UV} also satisfies $\theta = d\omega^V$. Define

$$c_{UV} = \exp(f_{UV}).$$

One easily verifies the properties $c_{UV}^{-1}=c_{VU}$ and $c_{VU}c_{UW}c_{WV}=id$. Thus the collection of complex valued 1-forms $\{\omega^U,\omega^V,\cdots,\}$ with transition functions given by $c_{UV}=\exp(f_{UV})$ for any and all complex functions on M define a connection on the complex line bundle $\{E,M,\mathbb{C},\pi\}$.

When $M=\mathbb{R}^4$ is Minkowski space we see that the electromagnetic filed 2-form $\mathscr{F}$ is closed and since $M=\mathbb{R}^4$ is simply connected is exact as well. Thus we see that there exists $\mathscr{A} = \phi dt + \frac{1}{c}\mathscr{A}_s$ such that $\mathscr{F}=d\mathscr{A}$ and hence that the collection of 1-forms

$$\omega^U = -\left(\frac{ie}{\hbar}\right)\mathscr{V}^U = -\left(\frac{ie}{\hbar}\right)\left(\phi\, dt + \frac{1}{c}\mathscr{A}_s\right) \tag{5.8}$$

define a connection of the complex line bundle with transition functions given by

$$c_{UV} = \exp\left(-\left(\frac{ie}{\hbar}\right) f_{UV}\right),$$

for each real valued function f_{UV} on $M=\mathbb{R}^4$. Note that the curvature of this line bundle is simply

$$\theta = -\left(\frac{ie}{\hbar}\right)\mathscr{F}.$$

and gives rise to the electromagentic field in $M=\mathbb{R}^4$.

In the previous chapter we saw that the motion of a classical particle in an electromagnetic field with potential $\mathscr{A}^U = \phi\, dt + \frac{1}{c}\mathscr{A}_s$ is given by the Hamilton's equations with Hamiltonian $H^e \triangleq H(t,q,p^e - e\mathbf{A}) - e\phi(t,q)$

$$\dot{q} = \frac{\partial H^e}{\partial p^e},$$
$$\dot{p}^e = -\frac{\partial H^e}{\partial q}.$$

Quantizing the classical Hamiltonian using $p_j^e = -i\hbar\partial_{q_j}$ we have the quantum Hamiltonian

$$H^e = \frac{1}{2m}\sum_j^3 \left(-i\hbar\partial_{q_j} - e\mathbf{A}_j\right)^2 + V - e\phi,$$

and hence the corresponding Schrödinger's equation

$$i\hbar\frac{\partial\psi}{\partial t} = \frac{-\hbar^2}{2m}\sum_j^3\left(\partial_{q_j} - \left(\frac{ie}{\hbar}\right)\mathbf{A}_j\right)^2\psi + V\psi - e\phi\psi.$$

Re-arranging we have

$$i\hbar\left(\frac{\partial}{\partial t} - \left(\frac{ie}{\hbar}\right)\phi\right)\psi = \frac{-\hbar^2}{2m}\sum_j^3\left(\partial_{q_j} - \left(\frac{ie}{\hbar}\right)\mathbf{A}_j\right)^2\psi + V\psi.$$

Hence we have the following:

The quantum description of the motion of the classical particle in an electromagnetic field is given by the equations

$$i\hbar\nabla_{\partial_0}\psi = \frac{-\hbar^2}{2m}\sum_j^3\nabla_{\partial_j}\nabla_{\partial_j}\psi + V\psi.$$

where the wave function ψ is a cross section of the complex line bundle $\{E,\mathbb{R}^4,\mathbb{C},\pi\}$ with transition functions $c_{UV} = \exp\left(-\left(\frac{ie}{\hbar}\right)f_{UV}\right)$ where f_{UV} is a smooth function on $\mathbb{R}^4$.

Consider a single global covering of $\mathbb{R}^4$. Then any differentiable function f_{UV} defines a bundle with charts $\{(\mathbb{R}^4,\phi_U),(\mathbb{R}^4,\phi_V)\}$ with the transition map $c_{UV} = \exp\left(-\left(\frac{ie}{\hbar}\right)f_{UV}\right)$. Thus we have the following theorem:

Theorem 5.1. *If ψ^U is a solution of the Schrödinger's equation with electromagnetic potential $\mathscr{A}^U = \left(\phi\,dt + \frac{1}{c}\mathscr{A}_s\right)$ then so is*

$$\exp\left(-\left(\frac{ie}{\hbar}\right)f(t,x)\right)\psi(t,x).$$

with the potential replaced by $\mathscr{A}^V = \mathscr{A}^U + df = \left(\phi\,dt + \frac{1}{c}\mathscr{A}_s + df\right)$.

Since $|\exp\left(-\left(\frac{ie}{\hbar}\right)f(t,x)\right)\psi(t,x)| = |\psi(t,x)|$ we see that the probability interpretation of the wave function is unaltered by a change of potential by adding df.

Chapter 6
Transport Phenomena

6.1 General Formulation

A time dependent k-form, ω_M, on a manifold M^n is viewed as a parameterized form on M^n. Alternatively it can be considered as a k-form on the manifold $\mathbb{R}\times M^n$, since any k-form, $\sigma(t,m)$, on $\mathbb{R}\times M^n$ can be uniquely decomposed as

$$\sigma = dt\wedge\alpha_M + \beta_M,$$

where both α_M,β_M satisfy $i_{\partial_t}\alpha_M = 0$ and $i_{\partial_t}\beta_M = 0$ and β_M is a k-form on M^n and α_M is a $k-1$-form on M^n. Forms that satisfy $i_{\partial_t}\omega = 0$ do not have a dt term and are called spatial forms. Thus any time dependent spatial k-form β_M and a $k-1$-from α_M on M^n define a k-form on $\mathbb{R}\times M^n$. Furthermore any time dependent vectorfield X_M on M^n can be identified with a vectorfield X on $\mathbb{R}\times M^n$ using

$$X = \partial_t + X_M.$$

It follows that the flow of X projected onto M^n is the flow of X_m. The d operator on $\mathbb{R}\times M^n$ can also be uniquely written down as

$$d = dt\wedge\frac{\partial}{\partial t} + \mathbf{d}_M$$

where $\mathbf{d}_M$ is the d operator on M^n. Then

$$\begin{aligned}
d\sigma &= -dt\wedge d\alpha_M + d\beta_M = dt\wedge\left(\frac{\partial\beta_M}{\partial t} - \mathbf{d}\alpha_M\right) + \mathbf{d}\beta_M,\\
i_X d\sigma &= -dt\wedge i_{X_M}\left(\frac{\partial\beta_M}{\partial t} - \mathbf{d}\alpha_M\right) + \left(\frac{\partial\beta_M}{\partial t} - \mathbf{d}\alpha_M + i_{X_M}\mathbf{d}\beta_M\right),\\
&= -dt\wedge i_{X_M}\left(\frac{\partial\beta_M}{\partial t} + i_{X_M}\mathbf{d}\beta_M - \mathbf{d}\alpha_M\right) + \left(\frac{\partial\beta_M}{\partial t} + i_{X_M}\mathbf{d}\beta_M - \mathbf{d}\alpha_M\right).
\end{aligned}$$

Thus

Lemma 6.1. $i_X d\sigma = 0$ *if and only if*

$$\left(\frac{\partial \beta_M}{\partial t}+i_{X_M}\mathbf{d}\beta_M-\mathbf{d}\alpha_M\right)=0. \tag{6.1}$$

and hence if and only if

$$\mathbf{d}\alpha_M=\frac{\partial \beta_M}{\partial t}+\mathscr{L}^M_{X_M}\beta_M-\mathbf{d}i_{X_M}\beta_M. \tag{6.2}$$

A pair (X,σ) that satisfy $i_X d\sigma=0$ has several interesting properties. Consider a $k+1$-dimensional compact set, S_0, in $\mathbb{R}\times M^n$ with boundary $C_0=\partial S_0$. Note that S_0 need not lie on a slice $t=$ constant. Then $C_0=\partial S_0$ is a k-dimensional cycle. Consider the compact set $S_\tau=\phi^X_\tau(S_0)$ and the cycle $\partial S_\tau=C_\tau=\phi^X_\tau(C_0)$. Then the $k+1$ dimensional space $\Sigma_t=\cup_{0\leq\tau\leq t}C_\tau$ is the surface of the tube of solutions of X that originate from S_0. Then $\partial\Sigma_t=C_0-C_t$ and thus the Stoke's theorem implies that

$$\int_{\Sigma_t}d\sigma=\int_{\partial\Sigma_t}\sigma=\int_{C_0}\sigma-\int_{C_t}\sigma.$$

Also note that since $\partial C_t=0$

$$\frac{d}{dt}\int_{C_t}\sigma=\int_{C_t}\mathscr{L}_X\sigma=\int_{C_t}(i_X d\sigma+di_X\sigma)=\int_{C_t}i_X d\sigma+\int_{\partial C_t}i_X\sigma=\int_{C_t}i_X d\sigma,$$

we see that $\int_{\Sigma_t}d\sigma=0$ for every Σ_t if and only if $i_X d\sigma=0$. Summarizing we have:

Theorem 6.1. Kelvin's Theorem: *For two k-cycles that bound the surface of a tube of solutions of X*

$$\int_{C_0}\sigma=\int_{C_t}\sigma.$$

if and only if $i_X d\sigma=0$.

Furthermore since $C_0=\partial S_0$ we have from Stoke's theorem:

Lemma 6.2. *For two $k+1$-surface that slice a tube of solutions of X*

$$\int_{S_0}d\sigma=\int_{S_t}d\sigma. \tag{6.3}$$

if and only if $i_X d\sigma=0$.

For a given σ consider the distribution $\mathscr{D}=\{\zeta : i_\zeta d\sigma=0 \quad \& \quad i_\zeta dt=0\}$. Since for any $\zeta,\eta\in\mathscr{D}$

$$\begin{aligned}&i_{[\zeta,\eta]}d\sigma=\mathscr{L}_\zeta i_\eta d\sigma-i_\zeta\mathscr{L}_\eta d\sigma=-i_\zeta\mathscr{L}_\eta d\sigma=-i_\zeta(i_\eta dd\sigma+di_\eta d\sigma)=0,\\&i_{[\zeta,\eta]}dt=0,\end{aligned}$$

we see that $\mathscr{D}$ is involutive. The integral surfaces of $\mathscr{D}$ lie entirely in constant t slices. Consider a vectorfield X on $\mathbb{R}\times M^n$ such that $i_X dt=1$ and $i_X d\sigma=0$. Thus $X\notin\mathscr{D}$. However $\mathscr{L}_X d\sigma=i_X dd\sigma+di_X d\sigma=0$ and $\mathscr{L}_X dt=0$ and hence $(\phi^X_t)^*d\sigma=\sigma$ and $(\phi^X_t)^*dt=dt$. Thus we see that $\phi^X_t(\mathscr{D})=\mathscr{D}$ and thus we have the following:

Lemma 6.3. *For a given σ the integral surfaces of the distribution $\mathscr{D}=\{\zeta : i_\zeta d\sigma = 0 \quad \& \quad i_\zeta dt = 0\}$ are invariant under the flow of any vectorfield X that satisfies $i_X d\sigma = 0$.*

Recall that

$$d\sigma = dt \wedge \left(\frac{\partial \beta_M}{\partial t} - \mathbf{d}\alpha_M\right) + \mathbf{d}\beta_M$$

Thus for a spatial vectorfield ζ_M we have

$$i_{\zeta_M} d\sigma = -dt \wedge i_{\zeta_M}\left(\frac{\partial \beta_M}{\partial t} - \mathbf{d}\alpha_M\right) + i_{\zeta_M}\mathbf{d}\beta_M.$$

Thus if $i_X d\sigma = 0$ then from (6.1) we see that

$$\begin{aligned} i_{\zeta_M} d\sigma &= -dt \wedge i_{\zeta_M}\left(\frac{\beta_M}{\partial t} - \frac{\partial \beta_M}{\partial t} - i_{X_M}\mathbf{d}\beta_M\right) + i_{\zeta_M}\mathbf{d}\beta_M \\ &= dt \wedge i_{\zeta_M} i_{X_M}\mathbf{d}\beta_M + i_{\zeta_M}\mathbf{d}\beta_M = -dt \wedge i_{X_M} i_{\zeta_M}\mathbf{d}\beta_M + i_{\zeta_M}\mathbf{d}\beta_M. \end{aligned}$$

Thus we have

Lemma 6.4. *When $i_X d\sigma = 0$ the condition $i_{\zeta_M} d\sigma = 0$ holds true if and only if $i_{\zeta_M}\mathbf{d}\beta_M = 0$. Thus when $i_X d\sigma = 0$ we see that $\mathscr{D} = \mathscr{D}_M(t) = \{\zeta_M \,:\, i_{\zeta_M}\mathbf{d}\beta_M = 0\}$.*

The distribution $\mathscr{D}_M(t)$ is a time varying distribution on M^n. Lemma 6.3 then implies that:

Theorem 6.2. Helmholtz Theorem: *When X satisfies $i_X d\sigma = 0$ the time varying integral surfaces, $S(t)$, of $\mathscr{D}_M(t) = \{\zeta_M \,:\, i_{\zeta_M}\mathbf{d}\beta_M = 0\}$ are invariant under the flow of X. That is if $S(0)$ is an integral surface of $\mathscr{D}_M(0)$ then $\phi_t^X(S(0))$ is an integral surface of $\mathscr{D}_M(t)$.*

Helmholtz theorem states, in terms of fluid mechanics, that the surfaces $S(t)$ remain frozen in the fluid.

The statement $i_X d\sigma = 0$ has many physical consequences when one chooses (X, σ) appropriately.

6.2 Application to Fluid Mechanics

Let $V_M(t,m)$ be the time dependent vectorfield on M^n that gives the velocity of the fluid particle at m at time t and $V = \partial_t + V_M$ its version on $\mathbb{R}\times M^n$. Let ν_M be the unique spatial 1-form given by

$$\nu_M = \langle\langle V_M, \cdot \rangle\rangle. \tag{6.4}$$

Define

$$\begin{aligned} \sigma_{\text{vol}} &= -dt \wedge i_{V_M}\mu_M + \mu_M, \\ \sigma_\rho &= \rho\,\sigma_{\text{vol}}, \end{aligned}$$

$$\sigma = -\mathscr{E}dt + \nu_M,$$

where μ_M is the volume form on M^n and ρ is the density of the fluid and $\mathscr{E}$ is a function that will be defined later. Note that for any V

$$i_V\sigma_{\text{vol}} = -i_{V_M}\mu_M + i_{V_M}\mu_M = 0.$$

Notice that since $i_V\sigma_{\text{vol}} = 0$ for any V we also get

$$\begin{aligned}
i_V d\sigma_\rho &= i_V d(\rho\sigma_{\text{vol}}) = i_V\left(d\rho\wedge\sigma_{\text{vol}} + \rho d\sigma_{\text{vol}}\right) = i_V\left(d\rho\wedge\sigma_{\text{vol}}\right) + \rho\, i_V d\sigma_{\text{vol}}\\
&= d\rho(X)\sigma_{\text{vol}} - d\rho\wedge(i_V\sigma_{\text{vol}}) + \rho\, i_V d\sigma_{\text{vol}}\\
&= d\rho(V)\sigma_{\text{vol}} + \rho\, i_V d\sigma_{\text{vol}},\\
i_V\sigma_\rho &= \rho\, i_V\sigma_{\text{vol}} = 0.
\end{aligned}$$

We will show that $i_{V_M}d\sigma_i = 0$ where σ_i is any one of the above forms results any various conservation laws in fluid mechanics.

6.2.1 Incompressible Flow

Since

$$\mathscr{L}^M_{V_M}\mu_M = i_{V_M}\mathbf{d}\mu_M + \mathbf{d}i_{V_M}\mu_M,$$

from (6.1) we see that $i_V d\sigma_{\text{vol}} = 0$ if and only if $\mathscr{L}^M_{V_M}\mu_M = 0$. Note that since for any V $i_V\sigma_{\text{vol}} = -i_{V_M}\mu_M + i_{V_M}\mu_M = 0$ we have that if volumes are conserved by the flow of V then

$$\begin{aligned}
\frac{d}{dt}\int_{B(t)}\mu_M &= \frac{d}{dt}\int_{(t,B(t))}\mu_M = \frac{d}{dt}\int_{(t,B(t))}\sigma_{\text{vol}} = \int_{(t,B(t))}\mathscr{L}_V\sigma_{\text{vol}}\\
&= \int_{(t,B(t))} i_V d\sigma_{\text{vol}} + d i_V\sigma_{\text{vol}} = \int_{(t,B(t))} i_V d\sigma_{\text{vol}} = 0,
\end{aligned}$$

for all volumes $B(t)$. Thus we have the following:

Theorem 6.3. Incompressible Flow: *Volumes in M^n are conserved along the flow of V if and only if $i_V d\sigma_{\text{vol}} = 0$ and hence if and only if $\mathscr{L}^M_{V_M}\mu_M = (\nabla^M V_M)\,\mu_M = 0$.*

In terms of fluids this says that the flow is incompressible if and only if the divergence of V_M vanishes. Since $\mathscr{L}^M_{V_M}\mu_M = \mathbf{d}i_{V_M}\mu_M = 0$ the converse to the Poincare Lemma implies that there exists a 1-form α_M such that $\mathbf{d}\alpha_M = i_{V_M}\mu_M$.

6.2.2 Conservation of Mass

Notice that since $i_V\sigma_{\text{vol}} = 0$, $i_V d\sigma_\rho = d\rho(V)\sigma_{\text{vol}} + \rho\, i_V d\sigma_{\text{vol}}$, and $i_V\sigma_\rho = \rho\, i_V\sigma_{\text{vol}} = 0$ for any V

Hence we have that

$$\begin{aligned}\frac{d}{dt}\int_{B(t)}\rho\mu_M &= \frac{d}{dt}\int_{(t,B(t))}\sigma_\rho = \int_{(t,B(t))}\mathscr{L}_V\sigma_\rho = \int_{(t,B(t))} i_V d\sigma_\rho + d i_V\sigma_\rho \\ &= \int_{(t,B(t))} i_V d\sigma_\rho\end{aligned}$$

for all volumes $B(t)$. From (6.1) we have that $i_V d\sigma_\rho = 0$ if and only if

$$\begin{aligned} 0 &= \left(\frac{\partial\rho}{\partial t}\mu_M + i_{V_M}\mathbf{d}\rho\mu_M + \mathbf{d}i_{V_M}\rho\mu_M\right), \\ &= \left(\frac{\partial\rho}{\partial t}\mu_M + \mathscr{L}^M_{V_M}\rho\mu_M\right), \\ &= \left(\frac{\partial\rho}{\partial t}\mu_M + \mathbf{d}\rho(V_M)\mu_M + \mathscr{L}^M_{V_M}\mu_M\right), \\ &= \left(\frac{\partial\rho}{\partial t} + \mathbf{d}\rho(V_M) + \nabla^M V_M\right)\mu_M. \end{aligned}$$

Thus we have the following:

Theorem 6.4. Liouville Theorem: *The mass is conserved along the flow of V if and only if*

$$i_V d\sigma_\rho = d\rho(V)\sigma_{\mathrm{vol}} + \rho\, i_V d\sigma_{\mathrm{vol}} = 0, \tag{6.5}$$

and hence if and only if

$$\frac{\partial\rho}{\partial t} + \mathbf{d}\rho(V_M) + \nabla^M V_M = 0. \tag{6.6}$$

When ρ is a probability measure the above equation is the Fokker-Plank equation.

Thus in addition if the volume is conserved then the mass is conserved if and only if

$$d\rho(V) = \frac{\partial\rho}{\partial t} + \mathbf{d}\rho(V_M) = 0. \tag{6.7}$$

6.2.3 Euler-Bernoulli Flow

Recall that ν_M is the unique spatial 1-form given by $\nu_M = \langle\langle V_M, \cdot\rangle\rangle$ and consider

$$\sigma = -\mathscr{E}dt + \nu_M.$$

Then from (6.1) we have that $i_V d\sigma = 0$ if and only if

$$\left(\frac{\partial\nu_M}{\partial t} + i_{V_M}\mathbf{d}\nu_M + \mathbf{d}\mathscr{E}\right) = 0. \tag{6.8}$$

One can show explicitly that if

$$\mathscr{E} = \frac{1}{2}\langle\langle V_M, V_M\rangle\rangle + P + \Phi = \frac{1}{2} i_{V_M}\nu_M + P + \Phi, \tag{6.9}$$

where $\Phi(t,m)$ is the volume force potential and $P(t,m)$ is the pressure force potential given by

$$dP = \frac{1}{\rho}dp \tag{6.10}$$

then by direct verification one sees that (6.8) is Euler's non-stationary fluid flow equation for an ideal fluid. A fluid that satisfies the condition that the pressure depends only in density is called a *Barotropic Fluid*. The expression (6.10) holds only for a barotropic fluid.

That is we have that the non-stationary fluid flow equation for an ideal barotropic fluid is given by $i_V d\sigma = 0$ or equivalently by (6.8). The function (6.9) is called the Bernoulli function. From lemma 6.3, noting that $\beta_M = \nu_M$ we have that the integral lines of the 1D-distribution $\mathscr{D} = \mathscr{D}_M(t) = \{\zeta_M : i_{\zeta_M}\mathbf{d}\nu_M = 0\}$ are invariant under the flow of the vectorfield X and remain frozen in the fluid. These lines are called *Vortex Lines* and this is the Helmholtz theorem.

Euler's equation (6.8) shows that along the flow

$$i_{V_M}\frac{\partial \nu_M}{\partial t} + i_{V_M}\mathbf{d}\mathscr{E} = i_{V_M}\frac{\partial \nu_M}{\partial t} + \mathbf{d}\mathscr{E}(V_M) = 0. \tag{6.11}$$

Since

$$\frac{\partial}{\partial t}\langle\langle V_M, V_M\rangle\rangle = \frac{\partial}{\partial t}\langle \nu_M, V_M\rangle = i_{V_M}\frac{\partial \nu_M}{\partial t} + i_{\frac{\partial V_M}{\partial t}}\nu_M = 2 i_{V_M}\frac{\partial \nu_M}{\partial t}.$$

Summarizing we have the following theorem

Theorem 6.5. Euler-Bernoulli Theorem: *Ideal barotropic fluid flow satisfies the condition $i_V d\sigma = 0$ or equivalently*

$$\left(\frac{\partial \nu_M}{\partial t} + i_{V_M}\mathbf{d}\nu_M + \mathbf{d}\mathscr{E}\right) = 0. \tag{6.12}$$

Furthermore along the flow of a fluid particle the Bernoulli's function satisfies

$$\mathbf{d}\mathscr{E}(V_M) = -\frac{1}{2}\frac{\partial}{\partial t}\left(||V_M||^2\right). \tag{6.13}$$

In particular for stationary flows $\mathbf{d}\mathscr{E}(V_M) = 0$ and the Bernoulli's function is constant along the flow lines.

For irrotational flow $\mathbf{d}\nu_M = 0$. For irrotational flow $\mathbf{d}\nu_M = 0$ and then Euler's equation (6.12) becomes

$$\frac{\partial \nu_M}{\partial t} + \mathbf{d}\mathscr{E} = 0.$$

Since $\mathbf{d}\nu_M = 0$ the converse to the Poincare lemma implies that there exists a function Ψ such that $\nu_M = \mathbf{d}\Psi$.

Lemma 6.5. *Thus for ideal, barotropic, irrotational flow Euler's equation becomes*

$$\mathbf{d}\left(\frac{\partial \Psi}{\partial t} + \mathscr{E}\right) = 0, \tag{6.14}$$

or equivalently that

$$\frac{\partial \Psi}{\partial t} + \frac{1}{2}||\nabla^M \Psi||^2 + P + \Phi = \gamma(t), \tag{6.15}$$

for some function of t given by $\gamma(t)$.

In addition if the flow is stationary then the Bernoulli's function is constant everywhere.

From theorem 6.3 for incompressible irrotational flow we have $\nabla^M \cdot \nabla^M \Psi = 0$.

From lemma 6.1 Kelvin's theorem follows: For two cycles that bound the surface of a tube of solutions of V

$$\int_{C_0} \sigma = \int_{C_t} \sigma.$$

Chapter 7
Elements of Differential Geometry

7.1 Vectors and Covectors

Let M^n be a smooth manifold and $\mathscr{U}_1, \mathscr{U}_2, \cdots$ be a collection of open sets on M^n such that it covers M^n. That is $M^n = \cup_{\alpha \in \Lambda} \mathscr{U}_\alpha$. Consider the 1-1 and onto map $\phi_\alpha : \mathscr{U}_\alpha \to \phi_\alpha(\mathscr{U}_\alpha) \subset \mathbb{R}^n$. The pair $\{\mathscr{U}_\alpha, \phi_{U_\alpha}\}$ is called a coordinate chart. For a point $p \in M^n$ denote $x_U \triangleq \phi_U(p)$. Where $x_U = (x_U^1, x_U^2, \cdots, x_U^n) \in \mathbb{R}^n$ is called the coordinates of the point p in the patch $\{\mathscr{U}, \phi_U\}$. If $\{\mathscr{V}, \phi_V\}$ is a different coordinate patch that contains p then $x_V \triangleq (x_V^1, x_V^2, \cdots, x_V^n) \in \mathbb{R}^n$ is called the coordinates of the point p in the patch $\{\mathscr{V}, \phi_V\}$. The two coordinate representations of the point p are thus related by $x_V = f_{VU}(x_U)$ where $f_{VU} = \phi_V \circ \phi_U^{-1}$.

Let $c : [-T, T] \to M^n$ be a curve on M_n such that $c(0) = p \in M^n$. The tangent to the curve $c(t)$ at $p \in M^n$ is defined to be $\mathrm{v}_U \triangleq \left.\frac{d}{dt}\right|_{t=0} \phi_U(c(t)) = \dot{x}_U$ in the coordinate patch $\{\mathscr{U}, \phi_U\}$. If $\{\mathscr{V}, \phi_V\}$ a different coordinate patch that contains p then $\mathrm{v}_V \triangleq \left.\frac{d}{dt}\right|_{t=0} \phi_V(c(t)) = \dot{x}_V$. Conventionally we will treat v_U and v_V as a column matrix. Thus using the chain rule we see that

$$\mathrm{v}_V = \left[\frac{\partial f_{VU}}{\partial x_U}\right] \mathrm{v}_U = \left[\frac{\partial x_V}{\partial x_U}\right] \mathrm{v}_U . \tag{7.1}$$

Notice that $\left[\frac{\partial x_U}{\partial x_V}\right] = \left[\frac{\partial x_V}{\partial x_U}\right]^{-1}$. A *tangent vector* at p is defined to be an assignment of an n-tuple v_U corresponding to each coordinate patch $\{\mathscr{U}, \phi_U\}$ such that for any two patches $\{\mathscr{U}, \phi_U\}$ and $\{\mathscr{V}, \phi_V\}$ the two corresponding assignment v_U and v_V are related by (7.1). Physically it represents the tangent to some curve, $c(t)$, that passes through p. Symbolically we will denote this by $\mathrm{v}_p = \dot{c}(0)$ and we will say that v_p has the representation v_U in the coordinate chart $\{\mathscr{U}, \phi_U\}$ and that two such representations are related by (7.1). Observe that if the same curve is parameterised by a different parameter τ then the tangent vector of the curve $c(\tau)$ at p is $\mathrm{v}'_p = \dot{c}(0)\frac{dt}{d\tau} = v_p \frac{dt}{d\tau}$ and thus we see that different parameterisations will only change the magnitude of the tangent vector. On the other hand a complete different curve through p will give tangent vector at p with a different direction. The space of all such tangent vectors at p is called the *tangent space* to M^n at p and will be denoted by T_pM^n. Notice that T_pM^n is an n-dimensional vector space.

Let $f : M^n \to \mathbb{R}$ be a function defined on M^n and v_p be a tangent vector at p. Let $c(t)$ be a curve on M^n such that $p = c(0)$ and $\mathrm{v}_p = \dot{c}(0)$. Then we define

$$\mathrm{v}_p(f) \triangleq \left.\frac{df(c(t))}{dt}\right|_{t=0}. \tag{7.2}$$

From a physical point of view this represents the rate of change of f as one moves along the curve $c(t)$ and thus we are motivated to call $\mathrm{v}_p(f)$ the *directional derivative* of f along the direction v_p. In a particular coordinate patch $\{\mathscr{U}, \phi_U\}$ evaluating (7.2) we have

$$\mathrm{v}_p(f) = \left[\frac{\partial f}{\partial x_U}\right]\mathrm{v}_U = \sum_{i=1}^{n} \frac{\partial f}{\partial x_U^i}\mathrm{v}_U^i. \tag{7.3}$$

Similarly in a different coordinate patch $\{\mathscr{V}, \phi_V\}$ we will have $\mathrm{v}_p(f) = \left[\frac{\partial f}{\partial x_V}\right]\mathrm{v}_V$. Since $f \circ \phi_U^{-1} = f \circ \phi_V^{-1}$ we see from the chain rule that

$$\left[\frac{\partial f}{\partial x_U}\right]\mathrm{v}_U = \left[\frac{\partial f}{\partial x_V}\right]\left[\frac{\partial x_V}{\partial x_U}\right]\mathrm{v}_U = \left[\frac{\partial f}{\partial x_V}\right]\mathrm{v}_V.$$

Thus we see that the value of $\mathrm{v}_p(f)$ is independent of the choice of the coordinates, as it should be for intrinsic quantities.

By this definition it can be seen that $\mathrm{v}_p(\alpha f + \beta g) = \alpha \mathrm{v}_p(f) + \beta \mathrm{v}_p(g)$ and $\mathrm{v}_p(fg) = \mathrm{v}_p(f)\,g + \mathrm{v}_p(g)\,f$ for any $\alpha, \beta \in \mathbb{R}$ and functions f, g on M^n. Thus each v_p can be thought of as a *derivation* (ie. a linear operator on functions on M^n that satisfies the Liebnitz rule). That is if $\mathscr{F}(M^n)$ is the space of all smooth functions on M^n then $\mathrm{v}_p : \mathscr{F}(M^n) \to \mathbb{R}$ is a derivation for each v_p. In the coordinate patch $\{\mathscr{U}, \phi_U\}$ we may thus write

$$\mathrm{v}_p = \sum_{i=1}^{n} \mathrm{v}_U^i \frac{\partial}{\partial x_U^i} = \left[\frac{\partial}{\partial x_U}\right]\mathrm{v}_U, \tag{7.4}$$

where $\frac{\partial}{\partial x_U^i} : \mathscr{F}(M^n) \to \mathbb{R}$ is defined simply to be operator that takes the function $f \in \mathscr{F}(M^n)$ to the value of $\frac{\partial f}{\partial x_U^i}(p)$ in the coordinate patch $\{\mathscr{U}, \phi_U\}$. Notice that if the curve $c_i(t)$ passing through p at $t = 0$ is such that its representation in the coordinate patch $\{\mathscr{U}, \phi_U\}$ takes the form $\phi_U(c_i(t)) = (0, 0, \cdots, x_U^i(t), \cdots, 0)$ then the representation of the tangent vector along the curve $c_i(t)$ that is denoted by $\mathrm{v}_p^i = \dot{c}_i(0)$ takes the form $\mathrm{v}_U^i = (0, 0, \cdots, \dot{x}_U^i(t), \cdots, 0)$ in the coordinate patch $\{\mathscr{U}, \phi_U\}$. Consider a parameterisation of the curve such that $\dot{x}_U^i(t) \equiv 1$ along the curve. Then $\mathrm{v}_U^i = (0, 0, 1, \cdots, 0)$ and we see that for any $f \in \mathscr{F}$ the quantity $v_p^i(f) = \frac{\partial f}{\partial x_U^i}(p)$. This quantity gives the rate of change of f as we move along the curve $c_i(t)$. Thus we can consider $\dfrac{\partial}{\partial x_U^i}$ to represent the directional derivative operator in the i^{th} coordinate direction.

The expression (7.4) also shows that any $v_p \in T_pM^n$ can be expressed as a linear combination of the set of differential operators $\dfrac{\partial}{\partial x_U} = \left\{\dfrac{\partial}{\partial x_U^1}, \dfrac{\partial}{\partial x_U^2}, \cdots, \dfrac{\partial}{\partial x_U^n}\right\}$. Therefore we may consider $\dfrac{\partial}{\partial x_U}$ as a basis for T_pM^n in the coordinate patch $\{\mathscr{U}, \phi_U\}$. Similarly $\dfrac{\partial}{\partial x_V}$ will be a basis for T_pM^n in the coordinate patch $\{\mathscr{V}, \phi_V\}$ and from (7.4) and the coordinate transformation property of a tangent vector given by (7.1) that the two representations of the basis of T_pM^n are related by

$$\left[\frac{\partial}{\partial x_V}\right] = \left[\frac{\partial}{\partial x_U}\right]\left[\frac{\partial x_U}{\partial x_V}\right]. \tag{7.5}$$

Let $\alpha_p : T_pM^n \to \mathbb{R}$ be a linear function (mostly referred to as a linear functional). Being a linear function α_p is uniquely determined by its values on basis vectors of T_pM^n. Thus let us consider a coordinate chart $\{\mathscr{U}, \phi_U\}$ and the representation of the basis of T_pM^n in it given by $\left[\dfrac{\partial}{\partial x_U}\right]$. Denote by $\alpha_i^U \triangleq \alpha\left(\dfrac{\partial}{\partial x_U^i}\right)$. The elements $\alpha^U = \left[\alpha_1^U, \alpha_2^U, \cdots, \alpha_n^U\right]$ uniquely determines the linear operator α_p since from linearity we have

$$\alpha_p(\mathrm{v}_p) = \alpha\left(\sum_{i=1}^n \mathrm{v}_U^i \frac{\partial}{\partial x_U^i}\right) = \sum_{i=1}^n \mathrm{v}_U^i \alpha\left(\frac{\partial}{\partial x_U^i}\right) = \sum_{i=1}^n \mathrm{v}_U^i \alpha_i^U = \alpha^U \mathrm{v}_U$$

We will say that α^U is the representation of α_p in the $\{\mathscr{U}, \phi_U\}$ coordinates. Similarly let α^V be the representation of α_p in the $\{\mathscr{V}, \phi_V\}$ coordinates. Since $\alpha(\mathrm{v}_p)$ should have the same value irrespective of the coordinates being used to evaluate $\alpha_p(\mathrm{v}_p)$ we should have

$$\alpha_p(\mathrm{v}_p) = \alpha^V \mathrm{v}_V = \alpha^U \mathrm{v}_U = \alpha^U \left[\frac{\partial x_U}{\partial x_V}\right] \mathrm{v}_V$$

and hence that the two representations of α_p in overlapping coordinate patches should be related by

$$\alpha^V = \alpha^U \left[\frac{\partial x_U}{\partial x_V}\right]. \tag{7.6}$$

In differential geometric terminology the linear functional α_p will be called a *co-vector* at p. It is easy to see that the set of all co-vectors at p is also an n-dimensional vector space. We will call this the *co-tangent* space, $T_p^*M^n$, at p. Let us try to find a representation for the basis of $T_p^*M^n$ in the $\{\mathscr{U}, \phi_U\}$ coordinates. We will define $dx_U^i : T_pM^n \to \mathbb{R}^n$ such that $dx_U^i\left(\dfrac{\partial}{\partial x_U^j}\right) = \delta_j^i$. Then we see that we may express α_p as a linear combination of the $dx_u = \{dx_U^1, dx_U^2, \cdots, dx_U^n\}$ that is we can write

$$\alpha_p = \sum_{i=1}^{n} \alpha_i^U dx_U^i = \alpha^U \, dx_U. \tag{7.7}$$

Once again as for tangent vectors we observe that $\alpha^U \, dx_U = \alpha^V \, dx_V$ and hence from the transformation rule (7.7) that

$$\alpha^V \, dx_V = \alpha^V \left[\frac{\partial x_V}{\partial x_U}\right] dx_U$$

and hence that

$$dx_V = \left[\frac{\partial x_V}{\partial x_U}\right] dx_U. \tag{7.8}$$

This is nothing but a re-statement of the chain rule for the coordinate transformation map.

The collection of all tangent planes to M^n will be referred to as the *tangent bundle*, TM^n, while the collection of all co-tangent planes to M^n will be referred to as the *co-tangent bundle*, T^*M^n. In a coordinate patch $\{\mathscr{U}, \phi_U\}$ an element $(p, \mathrm{v}_p) \in TM^n$ will have the representation (x_U, v_U) and an element $(p, \alpha_p) \in T^*M^n$ will have the representation (x_U, α^U). In each of these cases we can define a projection map $\pi : TM^n \to M^n$ such that $\pi(p, \mathrm{v}_p) = p$ and $\pi : T^*M^n \to M^n$ such that $\pi(p, \alpha_p) = p$.

A smooth assignment of a tangent vector at each point p is called a *vector field*, on M^n. Explicitly it is a map $X : M^n \to TM^n$ such that $\pi \circ X = id$. Similarly a smooth assignment of a cotangent vector at each point p is called a *1-form field* (co-vector filed), on M^n. Explicitly it is a map $\alpha : M^n \to T^*M^n$ such that $\pi \circ \alpha = id$. In a coordinate patch $\{\mathscr{U}, \phi_U\}$ on M^n we may write these as

$$X(p) = \sum_{i=1}^{n} X_U^i(p) \frac{\partial}{\partial x_U^i}, \tag{7.9}$$

$$\alpha(p) = \sum_{i=1}^{n} \alpha_i^U(p) dx_U^i. \tag{7.10}$$

Let X and Y be smooth vector fields on M^n and let ϕ_X^t be the flow of X. We define the Lie-derivative of Y in the direction of X by the relationship

$$\mathscr{L}_X Y \triangleq \left.\frac{d}{dt}\right|_{t=0} {\phi_X^t}^* Y(\phi_t(p)) \triangleq \lim_{t\to 0} \frac{T_{\phi_t(p)}\phi_X^{-t} Y(\phi_t(p)) - Y(p)}{t} = \lim_{t\to 0} \frac{Y(\phi_t(p)) - T_p\phi_X^t Y(p)}{t}. \tag{7.11}$$

7.2 Tensor Fields and Forms

Analogous to the definition of a vector field and a co-vector field we may define higher dimensional functional. For instance we may define the map $Q_p : T_pM^n \times T_pM^n \times \cdots \times T_pM^n \to \mathbb{R}$ such that Q_p is symmetric and multi-linear. If the number of products of T_pM^n in the above definition is r we will call it a symmetric covariant form of rank r. Using the linearity property, in a coordinate patch $\{\mathscr{U}, \phi_U\}$ on M^n, suppressing the indication of the coordinate chart, we may write this form as

$$Q_p = \sum_{\mathbf{J}} Q_{\mathbf{J}}\, dx^{j_1} \otimes dx^{j_2} \otimes \cdots \otimes dx^{j_r} \tag{7.12}$$

where $\mathbf{J} = (j_1, j_2, \cdots, j_r)$ with each $j_i \in (1, 2, \cdots, r)$ while

$$dx_U^{j_1} \otimes dx_U^{j_2} \otimes \cdots \otimes dx_U^{j_r} \left(\frac{\partial}{\partial x_U^{i_1}}, \frac{\partial}{\partial x_U^{i_2}}, \cdots, \frac{\partial}{\partial x_U^{i_r}} \right) = \begin{cases} 1 \text{ if } \mathbf{I} = \mathbf{J} \\ \\ 0 \text{ if } \mathbf{I} \neq \mathbf{J} \end{cases}$$

for $\mathbf{I} = (i_1, i_2, \cdots, i_r)$ with each $i_i \in (1, 2, \cdots, r)$. A smooth assignment of covariant tensors of the above form at each p of the manifold similar to that of a vector field will be called a covariant symmetric form field or simply a *symmetric r-tensor*. We will denote the space of all such tensor fields as $\otimes^r TM^n$. An important tensor field that we frequently come across in mechanics is the *symmetric* 2*-tensor* that takes the form

$$G = \sum_{i,j} G_{ij}\, dx^i \otimes dx^j.$$

Requirement of symmetry implies that $G_{ij} = G_{ji}$.

Another commonly encountered form is a *skew symmetric r-form*. The skew symmetry implies that $\alpha\left(\cdots, \mathrm{v}_i, \cdots, \mathrm{v}_j, \cdots\right) = -\alpha\left(\cdots, \mathrm{v}_j, \cdots, \mathrm{v}_i, \cdots\right)$. That is the permutation of any two of the arguments causes only a change in sign. Such a form in a coordinate patch $\{\mathscr{U}, \phi_U\}$ on M^n can be expressed as

$$\alpha_p = \sum_{\mathbf{J}} \alpha_{\mathbf{J}}\, dx^{j_1} \wedge dx^{j_2} \wedge \cdots \wedge dx^{j_r} \tag{7.13}$$

where $\mathbf{J} = (j_1, j_2, \cdots, j_r)$ is an ordered permutation of $(1, 2, \cdots, r)$ and if $\mathbf{I} = (i_1, i_2, \cdots, i_r)$ with each $i_i \in (1, 2, \cdots, r)$ then

$$dx_U^{j_1} \wedge dx_U^{j_2} \wedge \cdots \wedge dx_U^{j_r} \left(\frac{\partial}{\partial x_U^{i_1}}, \frac{\partial}{\partial x_U^{i_2}}, \cdots, \frac{\partial}{\partial x_U^{i_r}} \right) = \delta_{\mathbf{I}}^{\mathbf{J}}.$$

Here

$$\delta_{\mathbf{I}}^{\mathbf{J}} = \begin{cases} 1 & \text{if } \mathbf{I} \text{ is an even permutation of } \mathbf{J} \\ -1 & \text{if } \mathbf{I} \text{ is an odd permutation of } \mathbf{J} \\ 0 & \text{if } \mathbf{I} \text{ is not a permutation of } \mathbf{J} \end{cases}$$

A smooth assignment of such a skew symmetric r-ranked bilinear field to every point of the manifold will be called a smooth *r-form*. The space of all such forms will be denoted by $\wedge^r M^n$.

If $\alpha \in \wedge^p M^n$ and $\beta \in \wedge^q M^n$ then we may define the $(p+q)$-form $(\alpha \wedge \beta) \in \wedge^{(p+q)} M^n$ by the relationship

$$(\alpha \wedge \beta)(\mathrm{v_I}) = \sum_{\mathbf{J}} \sum_{\mathbf{K}} \delta_{\mathbf{I}}^{\mathbf{JK}} \alpha(\mathrm{v_J}) \beta(\mathrm{v_K}). \tag{7.14}$$

Here we use the convention that index sets are ordered and $\mathrm{v_I} = \mathrm{v}_{i_1}, \cdots, \mathrm{v}_{i_{p+q}}$.

It can be shown that there exists an unique exterior differentiation operator

$$d : \overset{p}{\wedge} M^n \to \overset{(p+1)}{\wedge} M^n$$

such that

(i) d is additive.
(ii) df is a usual differential of the function.
(iii) $d(\alpha^p \wedge \beta^q) = d\alpha^p \wedge \beta^q + (-1)^p \alpha^p \wedge d\beta^q$.
(iv) $d^2\alpha = 0$ for all forms.

In coordinates one finds that this is given by

$$d\alpha = \sum_{\mathbf{J}} d\alpha_{\mathbf{J}} \wedge dx^{j_1} \wedge dx^{j_2} \wedge \cdots \wedge dx^{j_r} \tag{7.15}$$

Similarly it can be shown that for $\mathrm{v} \in T_pM^n$ the map $i_{\mathrm{v}} : \overset{p}{\wedge} M^n \to \overset{(p-1)}{\wedge} M^n$ explicitly defined by

$$i_{\mathrm{v}} \alpha^q(\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_{q-1}) = \alpha^q(\mathrm{v}, \mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_{q-1}) \tag{7.16}$$

satisfies

$$i_{\mathrm{v}}(\alpha^p \wedge \beta^q) = i_{\mathrm{v}}\alpha^p \wedge \beta^q + (-1)^p \alpha^p \wedge i_{\mathrm{v}}\beta^q. \tag{7.17}$$

That is, it is an *anti-derivation*.

Let X be a smooth vector field on M^n and let ϕ_t be its flow. We define the Lie-derivative of a r-form by the relationship

$$\mathscr{L}_X \alpha = \left.\frac{d}{dt}\right|_{t=0} \phi_t^* \alpha = \lim_{t \to 0} \frac{\phi_t^* \alpha_{\phi_t(p)} - \alpha_p}{t} \tag{7.18}$$

One can show the following results

$$\mathscr{L}_X \circ d = d \circ \mathscr{L}_X, \tag{7.19}$$
$$\mathscr{L}_X = i_X \circ d + d \circ i_X, \tag{7.20}$$
$$\mathscr{L}_X \circ i_Y - i_Y \circ \mathscr{L}_X = i_{[X,Y]}. \tag{7.21}$$

7.3 Integration of Forms

Let α^p be a p-form and $(S, o) \in M^n$ be a compact oriented p-dimensional subset contained in some coordinate chart $\{\mathscr{U}, \phi_U\}$ on M^n. Then the integral of α^p over (S, o) is defined by

$$\int_{(S,o)} \alpha^p = \int_{(S,o)} \alpha^U dx_U^1 \wedge \cdots \wedge dx_U^p = o(S) \int_{\phi_U(S)} \alpha^U dx_U^1 \cdots dx_U^p, \tag{7.22}$$

where $o(S) = \pm 1$ where the value is 1 if the basis $\left\{ \frac{\partial}{\partial x_U^1}, \frac{\partial}{\partial x_U^2}, \cdots, \frac{\partial}{\partial x_U^p} \right\}$ is positively oriented and is -1 if otherwise.

We will often need the theorem:

Theorem 7.1. Stoke's Theorem: *Let $S^p \subset M^n$ be a compact subset with boundary ∂S^p and let ω^{p-1} be a smooth $p-1$-form on M^n. Then*

$$\int_{\partial S} \omega^{p-1} = \int_S d\omega^{p-1}. \tag{7.23}$$

A p-form ω is said to be *closed* if $d\omega = 0$. It is said to be *exact* if there exists a $(p-1)$ form α such that $\omega = d\alpha$.

If ω is a p-form and α is an exact p-form then we have the following consequences of the Stoke's theorem:

(i) If S is compact set without a boundary and ω is exact, then $\int_S \omega = 0$.
(ii) If S is a compact set with boundary and ω is closed, then $\int_S \omega = 0$.

Let $\Phi : M^n \to W^r$ be a map between two manifolds, $S \subset M^n$ is a p-dimensional compact subset and ω a p-form on W^r. Then

$$\int_{\Phi(S)} \omega = \int_S \Phi^* \omega. \tag{7.24}$$

Let X be a vector field on M^n, let ϕ_t be its flow, and ω be a fixed p-form (ie. with coefficients independent of time). Consider $S(t) = \phi_t(S)$ and then

$$I(t) = \int_{S(t)} \omega = \int_{\phi_t(S)} \omega = \int_S \phi_t^* \omega.$$

Thus from

$$I(t+h) - I(t) = \int_S \phi_t^* (\phi_h^* \omega - \omega).$$

it follows that

$$\begin{aligned} \frac{d}{dt} \int_{S(t)} \omega &= \int_{S(t)} \mathscr{L}_X \omega, \qquad (7.25) \\ &= \int_{S(t)} (d\, i_X \omega + i_X\, d\omega), \end{aligned}$$

$$= \int_{\partial S(t)} i_X \omega + \int_{S(t)} i_X \, d\omega. \tag{7.26}$$

In the case where the coefficients of α depend on time we may consider the manifold $\mathbb{R} \times M^n$. Then we may denote the differential operator, d, on $\mathbb{R} \times M^n$ as

$$d = dt \wedge \frac{\partial}{\partial t} + \mathbf{d},$$

where $\mathbf{d}$ is the differential operator on M^n. Consider the p-dimensional compact subset $W(t) = (t, S(t)) \in \mathbb{R} \times M^n$ that results due to a deformation of the set $W(0) = (0, S)$ by the flow ψ_t of the vector field $\chi = \dfrac{\partial}{\partial t} + X$. That is $W(t) = \psi_t(0, S) = (t, S(t))$. Now one can use (7.26) above to show that for a spatial ω (ie. a form that does not have a dt term)

$$\begin{aligned}
\frac{d}{dt}\int_{S(t)} \omega &= \frac{d}{dt}\int_{(t,S(t))} \omega = \int_{W(t)} \mathscr{L}_\chi \omega, \\
&= \int_{W(t)} (d\, i_\chi \omega + i_\chi \, d\omega) = \int_{\partial W(t)} i_\chi \omega + \int_{W(t)} i_\chi \, d\omega, \\
&= \int_{\partial S(t)} i_X \omega + \int_{W(t)} i_\chi \left(dt \wedge \frac{\partial \omega}{\partial t} + \mathbf{d}\omega \right), \\
&= \int_{\partial S(t)} i_X \omega + \int_{S(t)} \left(\frac{\partial \omega}{\partial t} + i_X \mathbf{d}\omega \right) = \int_{S(t)} \left(\frac{\partial \omega}{\partial t} + \mathscr{L}_X \omega \right).
\end{aligned} \tag{7.27}$$

Where we have used (7.20) in the last step.

If the integration is over a fixed set S then $X \equiv 0$ and then (7.27) becomes

$$\frac{d}{dt}\int_S \omega = \int_S \frac{\partial \omega}{\partial t}, \tag{7.28}$$

Definition 7.1. We will say that the *first Betti number* of M^n is zero if every closed curve is the boundary of some compact surface in M^n.

Consider a closed 1-form α on M^n. Let $y \in M^n$ be any point on M^n. Define $f_{C_1}(x) = \int_{C_1} \alpha$, where C_1 is some curve that joins y to $x \in M^n$. If C_2 is some other curve that joins y to $x \in M^n$ let $f_{C_2}(x) = \int_{C_2} \alpha$. Consider the closed curve $C_y = C_1 - C_2$. That is the one composed of moving along C_1 from y to x and then coming back along C_2 from x to y. Then we see that

$$\int_{C_y} \alpha = \int_{C_1} \alpha - \int_{C_2} \alpha = f_{C_1}(x) - f_{C_2}(x).$$

If the first Betti number of M^n is zero then there exists a compact surface S such that $\partial S = C_y$. Thus from the Stoke's theorem we have

$$\int_{C_y} \alpha = \int_{\partial S} \alpha = \int_S d\alpha = 0.$$

Which says that $f_{C_1}(x) = f_{C_2}(x) = f_y(x)$ and hence that $f_y(x) = \int_{C_y} \alpha$ does not depend on the curve that joins y to x.

Consider a curve $x(t)$ such that $x = x(0)$ and $\dot{x}(0) = v$. Let X be a vector field defined around C_y such that $X(x) = v$ and $X(y) = 0$ let ϕ_t be its flow. Then

$$
\begin{aligned}
df_y(v) &= \left.\frac{d}{dt}\right|_{t=0} f_y(x(t)) = \left.\frac{d}{dt}\right|_{t=0} \int_{\phi_t(C_y)} \alpha = \int_{C_y} \mathscr{L}_X \alpha \\
&= \int_{C_y} d i_X \alpha = \int_{\partial C_y} i_X \alpha = i_{X(x)}\alpha - i_{X(y)}\alpha, \\
&= i_v \alpha = \alpha(v).
\end{aligned}
$$

Hence we see that $\alpha = df_y$ and hence α is exact. Observe that this result is also true if α is such that $\int_C \alpha = 0$ for any closed curve on M^n. Consider $f_z(x) = \int_{C_z} \alpha$ where $z \in M^n$ is a different point on M^n and C_z is a curve that joins z to x. Then we see that $f_z(x) = f_y(x) + c$ where $c = \int_{C_{zy}} \alpha$ is a constant. Thus there exists a function $f : M^n \to \mathbb{R}$ that is unique unto an additive constant such that $\alpha = df$. We summarise these results in the following theorem:

Theorem 7.2. *If α is a 1-form on M^n there exists a function $f : M^n \to \mathbb{R}$ that is unique upto an additive constant such that $\alpha = df$ (that is it is exact) if either of the following conditions are true:*

(i) The integral of α over any closed curve on M^n is zero.
(ii) The first Betti number of M^n is zero.

Let $\mathscr{V}$ be a 1-form and $\mathbf{V}$ be the associated vector such that $\mathscr{V} = \langle\langle \mathbf{V}, \cdot \rangle\rangle$. The $(n-1)$-form $*\mathscr{V}$ is defined by the relationship

$$*\mathscr{V} \triangleq i_{\mathbf{V}} \mathrm{vol},$$

while the divergence of a vector $\mathbf{V}$ is defined to be the n-form

$$\mathrm{div}(\mathbf{V}) \triangleq \mathscr{L}_{\mathbf{V}} \mathrm{vol}.$$

For a n-form α the 0-form $*\alpha$ is defined such that $\alpha = (*\alpha)\mathrm{vol}$[3].

From the general form of the Stoke's theorem we have

Theorem 7.3. Divergence Theorem:

$$\int_\Omega d * \mathscr{V} = \int_{\partial\Omega} * \mathscr{V}, \tag{7.29}$$

$$\int_\Omega \mathrm{div}(\mathbf{V}) \mathrm{vol} = \int_{\partial\Omega} i_{\mathbf{V}} \mathrm{vol}. \tag{7.30}$$

Consider the vector that is expressed in spherical polar coordinates by the expression

$$\mathbf{V} = \frac{1}{4\pi r^2}\partial_r. \tag{7.31}$$

Then

$$\mathscr{V} = \frac{1}{4\pi r^2}dr,$$
$$*\mathscr{V} = i_{\mathbf{V}}\mathrm{vol}^3 = \frac{1}{4\pi}\sin\varphi\, d\varphi\wedge d\theta$$

Hence

$$\int_\Omega d*\mathscr{V} = \int_{\partial\Omega} *\mathscr{V} = \frac{1}{4\pi}\int_{\partial\Omega}\sin\varphi\, d\varphi\wedge d\theta = 1.$$

Thus we see that the 1-form $\mathscr{V}$ that corresponds to the radial vector (7.31) is a solution to the Poisson equation

$$d*\mathscr{V} = \delta(p)\mathrm{vol}. \tag{7.32}$$

In fact it can be shown that this is the unique solution to the Poisson equation (7.32).

Note that from (7.30) the Poisson equation (7.32) then takes the form

$$*d*\mathscr{V} = \delta(p), \tag{7.33}$$
$$\mathrm{div}(\mathbf{V}) = \delta(p). \tag{7.34}$$

Thus the radial vector (7.31) is the unique solution to the Poisson equation (7.34). In Euclidean coordinates the radial vector (7.31) is

$$\mathbf{V} = \frac{1}{4\pi}\frac{\mathbf{r}}{||\mathbf{r}||^3}.$$

By shifting the origin to $\mathbf{r}'$ we have that the solution to the equation

$$*d*\mathscr{V} = \delta(\mathbf{r}-\mathbf{r}'), \tag{7.35}$$
$$\mathrm{div}(\mathbf{V}) = \delta(\mathbf{r}-\mathbf{r}') \tag{7.36}$$

is given by

$$\mathbf{V}_{\mathbf{r}'} = \frac{1}{4\pi}\frac{(\mathbf{r}-\mathbf{r}')}{||\mathbf{r}-\mathbf{r}'||^3}. \tag{7.37}$$

Let $\mathscr{V}_{\mathbf{r}'} = \langle\langle \mathbf{V}_{\mathbf{r}'}, \cdot\rangle\rangle$.

Let $f(\mathbf{r}')$ be a function with compact support Ω. Let us find the solution to the equation

$$*d*\mathscr{V} = \int_\Omega f(\mathbf{r}')\mathrm{vol}^3_{\mathbf{r}'}. \tag{7.38}$$

From linearity it follows that the solution to (7.38) is given by

$$\mathbf{V}(\mathbf{r}) = \frac{1}{4\pi}\int_\Omega \frac{f(\mathbf{r}')(\mathbf{r}-\mathbf{r}')}{||\mathbf{r}-\mathbf{r}'||^3}\mathrm{vol}. \tag{7.39}$$

Let ϕ be a function then we have the following:

Theorem 7.4. Gauss's Theorem: *The solution to the Poisson equation*

$$\nabla^2 \phi = \int_\Omega f(\mathbf{r}')\mathrm{vol}, \tag{7.40}$$

where $\Omega \subset \mathbb{R}^3$ *is compact, is given by*

$$\phi = -\frac{1}{4\pi}\int_\Omega \frac{f(\mathbf{r}')}{||\mathbf{r}-\mathbf{r}'||}\mathrm{vol}. \tag{7.41}$$

Consider the equation

$$\nabla^2 \mathbf{A} = -\int_\Omega \mathbf{f}(\mathbf{r}')\mathrm{vol}.$$

In Euclidean coordinates this equation boils down to three independent equations of the form of the Poisson's equation. Thus we have in Euclidean coordinates that the solution to the vector Poisson equation is given by,

$$\mathbf{A} = \int_\Omega \frac{\mathbf{f}(\mathbf{r}')}{||\mathbf{r}-\mathbf{r}'||}\mathrm{vol}, \tag{7.42}$$

7.4 Example: Minkowski Space

Minkowski space is defined to be $\mathbb{R}^4$ equipped with the Pseudo-Riemannian metric (1.9). A general 1-form, 2-form and 3-form in Minkowski space can be expressed in a Lorentz coordinate system $e = (t, \mathbf{x})$ as

$$\mathscr{V} = \phi dt + \frac{1}{c}\mathscr{A}_s, \tag{7.43}$$

$$\mathscr{F} = \mathscr{E}_s \wedge dt + \frac{1}{c}\mathscr{B}_s, \tag{7.44}$$

$$\mathscr{J} = \mathscr{D}_s - \mathscr{J}_s \wedge dt, \tag{7.45}$$

where

$$\begin{aligned}
\mathscr{E}_s &= E_1 dx^1 + E_2 dx^2 + E_3 dx^3,\\
\mathscr{A}_s &= A_1 dx^1 + A_2 dx^2 + A_3 dx^3,\\
\mathscr{B}_s &= B_1 dx^2 \wedge dx^3 + B_2 dx^3 \wedge dx^1 + B_3 dx^1 \wedge dx^2,\\
\mathscr{J}_s &= J_1 dx^2 \wedge dx^3 + J_2 dx^3 \wedge dx^1 + J_3 dx^1 \wedge dx^2,\\
\mathscr{D}_s &= \rho^e dx^1 \wedge dx^2 \wedge dx^3
\end{aligned}$$

are forms that do not contain dt terms[1]. Observe that these forms $\mathscr{E}_s, \mathscr{A}_s, \mathscr{B}_s, \mathscr{J}_s, \mathscr{D}_s$, are not Lorentz invariant. Meaning that when one express them in a general different coordinate system they may contain dt terms. On the other hand $\mathscr{V}$, $\mathscr{F}$ being true general forms they are Lorentz invariant meaning if $(t', \mathbf{x}')$ is a different Lorentz coordinate system then they can still be expressed as $\mathscr{V} = \phi' dt' + \mathscr{A}_s'$, $\mathscr{F} = \mathscr{E}_s' \wedge dt' + \mathscr{B}_s'$, $\mathscr{J} = \mathscr{D}_s' - \mathscr{J}_s' \wedge dt'$ where as $\mathscr{E}_s', \mathscr{A}_s', \mathscr{B}_s', \mathscr{J}_s', \mathscr{D}_s'$, may contain dt' terms.

Also denote by $\mathbf{E} = (\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3) \in \mathbb{R}^3$, $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3) \in \mathbb{R}^3$, $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3) \in \mathbb{R}^3$, $\mathbf{J} = (\mathbf{J}_1, \mathbf{J}_2, \mathbf{J}_3) \in \mathbb{R}^3$ vector fields on $\mathbb{R}^3$. Consider a world line $(t, \mathbf{x}(t))$ of a particle system in the e co-ordinates and let $\mathbf{v} = \dot{\mathbf{x}}(t) \in \mathbb{R}^3$. Then the intrinsic tangent vector at a point $(t, \mathbf{x}(t))$ is given by $U = \gamma(1, \mathbf{v})$ where $\gamma = \dfrac{dt}{d\tau} = (1 - ||\mathbf{v}||^2/c^2)$. One can show that

$$i_U \mathscr{F} = \gamma \mathbf{E}\cdot\mathbf{v}\, dt - \gamma\left(\left(E_1 + \frac{1}{c}(\mathbf{v}\times\mathbf{B})_1\right)dx^1 + \left(E_2 + \frac{1}{c}(\mathbf{v}\times\mathbf{B})_2\right)dx^2 + \left(E_3 + \frac{1}{c}(\mathbf{v}\times\mathbf{B})_3\right)dx^3\right),$$

Let F be the 4-vector corresponding to this 1-form. That is $F^\flat = \frac{1}{c^2} i_U \mathscr{F}$. It is explicitly given by the F that satisfies

$$ds^2(F, V) = \frac{1}{c^2} i_U \mathscr{F}(V),$$

for all 4-vectors V. Thus we see that

$$F = \gamma \frac{\mathbf{E}\cdot\mathbf{v}}{c^2}\frac{\partial}{\partial t} + \gamma\left(\left(E_1 + \frac{1}{c}(\mathbf{v}\times\mathbf{B})_1\right)\frac{\partial}{\partial x_1} + \left(E_2 + \frac{1}{c}(\mathbf{v}\times\mathbf{B})_2\right)\frac{\partial}{\partial x^2} + \left(E_3 + \frac{1}{c}(\mathbf{v}\times\mathbf{B})_3\right)\frac{\partial}{\partial x^3}\right) \tag{7.46}$$

and in component wise

$$F = \gamma\left(\frac{\mathbf{E}\cdot\mathbf{v}}{c^2}, \mathbf{E} + \frac{1}{c}(\mathbf{v}\times\mathbf{B})\right). \tag{7.47}$$

Thus we see that

$$\mathbf{f} = \left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right). \tag{7.48}$$

corresponds to a classical force acting on the particle P as observed by the inertial observer $\mathbf{e}$.

Define $\mathbf{d}_s = dx^1 \wedge \dfrac{\partial}{\partial x^1} + dx^2 \wedge \dfrac{\partial}{\partial x^2} + dx^3 \wedge \dfrac{\partial}{\partial x^3}$ then in Minkowski space $d = dt \wedge \dfrac{\partial}{\partial t} + \mathbf{d}_s$. Then

$$d\mathscr{F} = \mathbf{d}_s \mathscr{E}_s \wedge dt + \frac{1}{c}\frac{\partial \mathscr{B}_s}{\partial t} \wedge dt + \frac{1}{c}\mathbf{d}_s \mathscr{B}_s = \left(\mathbf{d}_s \mathscr{E}_s + \frac{1}{c}\frac{\partial \mathscr{B}_s}{\partial t}\right) \wedge dt + \frac{1}{c}\mathbf{d}_s \mathscr{B}_s. \tag{7.49}$$

Let $*\mathscr{F}$ be the Hodge star dual of $\mathscr{F}$. That is the $(4-2)$-form defined by

$$\mathscr{A} \wedge *\mathscr{F} = c\,\langle \mathscr{A}, \mathscr{F} \rangle\, \mathrm{vol}, \tag{7.50}$$

for all 2-forms $\mathscr{A}$ where

[1] The factor $\frac{1}{c}$ has been introduced in front of $\mathscr{B}_s$ for notational convenience later on.

$$\mathrm{vol} = \frac{1}{c^3} dt \wedge dx^1 \wedge dx^2 \wedge dx^3$$

and $< \mathscr{A}, \mathscr{F} >= \mathscr{A}_{j_i j_k} \mathscr{F}^{j_i j_k}$ where $\mathscr{F}^{j_i j_k} = g^{j_1 r_1} g^{j_2 r_2} \mathscr{F}_{r_1 r_2}$. Then

$$\mathscr{A} \wedge *(dt \wedge dx_i) = c \, \langle \mathscr{A}, (dt \wedge dx_i) \rangle \, \mathrm{vol}$$

For any two form $\mathscr{A}$ the right hand side is equal to

$$c \, \langle \mathscr{A}, (dt \wedge dx_i) \rangle \, \mathrm{vol} = c \, \mathscr{A}_{0i} g^{00} g^{ii} \mathrm{vol} = -c^3 \, \mathscr{A}_{0i} \mathrm{vol}.$$

Thus we have that

$$*(dx_i \wedge dt) = - * (dt \wedge dx_i) = dx_j \wedge dx_k = \star dx_i$$

where i, j, k is a cyclic permutation of $1, 2, 3$. Similarly

$$\mathscr{A} \wedge *(dx_i \wedge dx_j) = c \left\langle \mathscr{A}, (dx_i \wedge dx_j) \right\rangle \mathrm{vol} = c \, \mathscr{A}_{ij} g^{ii} g^{jj} \mathrm{vol} = c^5 \, \mathscr{A}_{ij} \mathrm{vol},$$

and hence

$$*(dx_i \wedge dx_j) = c \, dt \wedge dx_k = -c \, dx_k \wedge dt = -c \star (dx_i \wedge dx_j) \wedge dt.$$

Thus we see that

$$*\mathscr{F} = *(\mathscr{E}_s \wedge dt) + \frac{1}{c} * \mathscr{B}_s = \star \mathscr{E}_s - c \star \mathscr{B}_s \wedge dt, \tag{7.51}$$

where

$$\begin{aligned} \star \mathscr{E}_s &= E_1 dx^2 \wedge dx^3 + E_2 dx^3 \wedge dx^1 + E_3 dx^1 \wedge dx^2, \\ \star \mathscr{B}_s &= B_1 dx^1 + B_2 dx^2 + B_3 dx^3. \end{aligned}$$

We then also see that

$$d * \mathscr{F} = d_s \star \mathscr{E}_s + \left(\frac{\partial \star \mathscr{E}_s}{\partial t} - c d_s \star \mathscr{B}_s \right) \wedge dt, \tag{7.52}$$

Also note that

$$\mathscr{F} \wedge \mathscr{F} = -c^2 \mathbf{E} \cdot \mathbf{B} \, \mathrm{vol}, \tag{7.53}$$

and

$$\mathscr{F} \wedge *\mathscr{F} = c^3 \left(||\mathbf{B}||^2 - ||\mathbf{E}||^2 \right) \mathrm{vol}. \tag{7.54}$$

If $(t', \mathbf{x}')$ is a different Lorentz coordinate system then one can show that the vol form in these coordinates is also given by $(1/c^3) \, dt' \wedge dx'^1 \wedge dx'^2 \wedge dx'^3$. Thus we see that although $\mathscr{E}_s, \mathscr{A}_s, \mathscr{B}_s, \mathscr{J}_s, \mathscr{D}_s, \mathbf{v}, \mathbf{f}, \mathbf{E}, \mathbf{B}, \mathbf{A}, \mathbf{J}, \mathbf{D}$, are not Lorentz invariant quantities, from (7.50) and (7.53), we see that $\left(||\mathbf{B}||^2 - ||\mathbf{E}||^2 \right)$ and $\mathbf{E} \cdot \mathbf{B}$ are Lorentz invariant quantities since $\mathscr{F}$ is an intrinsic quantity that has a meaning independent of coordinates.

Consider the case where the 2-form $\mathscr{F}$ is closed. That is $d\mathscr{F}=0$. Then the Poincare lemma implies that $\mathscr{F}$ is exact. That is there exists some local 1-form $\mathscr{V}$ such that $\mathscr{F}=d\mathscr{V}$. Thus

$$\mathscr{F}=\left(\mathbf{d}_s\phi-\frac{1}{c}\frac{\partial\mathscr{A}_s}{\partial t}\right)\wedge dt+\frac{1}{c}\mathbf{d}_s\mathscr{A}_s, \tag{7.55}$$

$$*\mathscr{F}=\star\left(\mathbf{d}_s\phi-\frac{1}{c}\frac{\partial\mathscr{A}_s}{\partial t}\right)-c\star\mathbf{d}_s\mathscr{A}_s\wedge dt, \tag{7.56}$$

$$d*\mathscr{F}=\left(\mathbf{d}_s\star\mathbf{d}_s\phi-\frac{1}{c}\mathbf{d}_s\star\frac{\partial\mathscr{A}_s}{\partial t}\right)+\left(\star\frac{\partial\mathbf{d}_s\phi}{\partial t}-\frac{1}{c}\star\frac{\partial^2\mathscr{A}_s}{\partial t^2}-c\mathbf{d}_s\star\mathbf{d}_s\mathscr{A}_s\right)\wedge dt. \tag{7.57}$$

7.4.1 Integration

Let V^p be a fixed p-dimensional compact set in $\mathbb{R}^3$ with ∂V being its boundary. Then the set $W^p(t)=(t,V^p(t))$ is a p-dimensional compact set in Minkowski space-time. Consider the vector field $X=(1,\mathbf{v})$ and let its flow be $\psi_t(t',\mathbf{x}')$. We see that $W^p(t)=\psi_t(0,V^p)$. Let α be a p-form. Then

$$\begin{aligned}
\frac{d}{dt}\int_{W^p(t)}\alpha &= \int_{W^p(0)}\frac{d}{dt}\psi_t{}^*\alpha=\int_{W^p(0)}\psi_t{}^*\mathscr{L}_X\alpha=\int_{W^p(t)}\mathscr{L}_X\alpha=\int_{W^p(t)}(i_Xd\alpha+di_X\alpha)\\
&=\int_{W(t)}i_Xd\alpha+\int_{\partial W^p(t)}i_X\alpha\\
&=\int_{W^p(t)}i_X\left(dt\wedge\frac{\partial\alpha}{\partial t}+\mathbf{d}_s\alpha\right)+\int_{\partial W^p(t)}i_X\alpha.
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}\int_{W^2(t)}\mathscr{F} &= \int_{W^2(t)}i_X\left(dt\wedge\frac{\partial\mathscr{F}}{\partial t}+\mathbf{d}_s\mathscr{F}\right)+\int_{\partial W^2(t)}i_X\mathscr{F}\\
&=\int_{W^2(t)}i_X\left(\frac{1}{c}dt\wedge\frac{\partial\mathscr{B}_s}{\partial t}+dt\wedge\mathbf{d}_s\mathscr{E}_s\right)+\int_{\partial W^2(t)}i_X\left(\mathscr{E}_s\wedge dt+\frac{1}{c}\mathscr{B}_s\right),\\
&=\int_{W^2(t)}\left(\frac{1}{c}\frac{\partial\mathscr{B}_s}{\partial t}+\mathbf{d}_s\mathscr{E}_s\right)+\int_{\partial W^2(t)}i_X\left(\mathscr{E}_s\wedge dt+\frac{1}{c}\mathscr{B}_s\right)\\
&=\int_{\partial W^2(t)}i_X\left(\mathscr{E}_s\wedge dt+\frac{1}{c}\mathscr{B}_s\right)=\int_{\partial W^2(t)}-\mathscr{E}_s+\frac{1}{c}i_{\mathbf{v}}\mathscr{B}_s\\
&=\int_{\partial V^2(t)}-\mathscr{E}_s+\frac{1}{c}i_{\mathbf{v}}\mathscr{B}_s
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}\int_{W^3(t)}\mathscr{J} &= \int_{W^3(t)}i_X\left(dt\wedge\frac{\partial\mathscr{J}}{\partial t}+\mathbf{d}_s\mathscr{J}\right)+\int_{\partial W^3(t)}i_X\mathscr{J}\\
&=\int_{W^3(t)}i_X\left(dt\wedge\frac{\partial\mathscr{D}_s}{\partial t}+dt\wedge\mathbf{d}_s\mathscr{J}_s\right)+\int_{\partial W^3(t)}i_X\left(\mathscr{D}_s-\mathscr{J}_s\wedge dt\right),
\end{aligned}$$

$$
\begin{aligned}
&= \int_{W^3(t)} \left(\frac{\partial \mathscr{D}_s}{\partial t} + \mathbf{d}_s \mathscr{J}_s \right) + \int_{\partial W^3(t)} (i_{\mathbf{v}} \mathscr{D}_s - \mathscr{J}_s) \\
&= \int_{W^3(t)} \left(\frac{\partial \mathscr{D}_s}{\partial t} + \mathbf{d}_s \mathscr{J}_s \right) + \int_{W^3(t)} (d i_{\mathbf{v}} \mathscr{D}_s - d \mathscr{J}_s) \\
&= \int_{W^3(t)} \left(\frac{\partial \mathscr{D}_s}{\partial t} + \mathbf{d}_s i_{\mathbf{v}} \mathscr{D}_s \right) \\
&= \int_{V^3(t)} \left(\frac{\partial \mathscr{D}_s}{\partial t} + \mathbf{d}_s i_{\mathbf{v}} \mathscr{D}_s \right).
\end{aligned}
$$

7.5 Vector Calculus in $\mathbb{R}^3$

Let $\{x\}$ and $\{y\}$ be two coordinate systems. Then $dy = \dfrac{\partial y}{\partial x} dx$. Since $dy(\partial_y) = dx(\partial_x)$ we have $\partial_y \dfrac{\partial y}{\partial x} = \partial_x$. Thus $\alpha^y dy = \alpha^y \dfrac{\partial y}{\partial x} dx = \alpha^x dx$ and $\partial_x V^x = \partial_y \dfrac{\partial y}{\partial x} V^x = \partial_y V^y$ which gives $\alpha^x = \alpha^y \dfrac{\partial y}{\partial x}$, $V^y = \dfrac{\partial y}{\partial x} V^x$. Let $\mathbf{V} = v^i \dfrac{\partial}{\partial x_i}$, $\alpha = \alpha_i dx_i$.

Associate with $\mathbf{V}$ the 1-form $\mathscr{V}^1$ by the relationship

$$\langle\langle \mathbf{V}, \cdot \rangle\rangle \triangleq \mathscr{V}^1,$$

and associate with $\mathbf{V}$ the 2-form $*\mathscr{V}$ by the relationship

$$*\mathscr{V} \triangleq i_{\mathbf{V}} \mathrm{vol},$$

Let $\mathscr{B}^2$ be a two form. Then we define the 1-form $*\mathscr{B}$ associated to $\mathscr{B}^2$ by the relationship

$$
\begin{aligned}
i_{\mathbf{B}} \mathrm{vol} &= \mathscr{B}^2 \\
*\mathscr{B} &\triangleq \langle\langle \mathbf{B}, \cdot \rangle\rangle.
\end{aligned}
$$

We also have $i_{\mathbf{B}} \mathrm{vol} = **\mathscr{B}$. Similarly since $i_{\mathbf{V}} \mathrm{vol} = *\mathscr{V}$ where $\mathscr{V}^1 = \langle\langle \mathbf{V}, \cdot \rangle\rangle$ we have $\mathscr{V}^1 = **\mathscr{V}$.

That is

$$
\begin{aligned}
i_{\mathbf{B}} \mathrm{vol} &= \mathscr{B}, \\
*\mathscr{B} &= *i_{\mathbf{B}} \mathrm{vol} = \langle\langle \mathbf{B}, \cdot \rangle\rangle \\
**\mathscr{B} &= **i_{\mathbf{B}} \mathrm{vol} = \mathscr{B}^2, \\
**\mathscr{V} &= *i_{\mathbf{V}} \mathrm{vol} = \mathscr{V}^1.
\end{aligned}
$$

Define $\mathrm{div}(\mathbf{V}) = \nabla \cdot \mathbf{V}$ by the relationship

$$\mathrm{div}(\mathbf{V}) = \nabla \cdot \mathbf{V} = *d*\mathscr{V}.$$

We also define $\nabla(\nabla \mathbf{V})$ by the relationship

$$d * d * \mathscr{V} = \langle\langle \nabla(\nabla \mathbf{V}), \cdot \rangle\rangle.$$

Define $\mathrm{curl}(\mathbf{V}) = \nabla \times \mathbf{V}$ by the relationship

$$i_{\nabla\times\mathbf{V}}\mathrm{vol} = d\mathscr{V}.$$

Then

$$\begin{aligned} *i_{\nabla\times\mathbf{V}}\mathrm{vol} &= *d\mathscr{V} = \langle\langle \nabla \times \mathbf{V}, \cdot \rangle\rangle \\ d * i_{\nabla\times\mathbf{V}}\mathrm{vol} &= d * d\mathscr{V} = i_{\nabla\times\nabla\times\mathbf{V}}\mathrm{vol}. \end{aligned}$$

The Laplacian operator is defined to be

$$\Delta \triangleq *d*d - d*d*.$$

Let f be a function then the vector $\mathrm{grad} f = \nabla f$ is defined to be the vector associated with the 1-form df. That is,

$$\langle\langle \nabla f, \cdot \rangle\rangle = df.$$

Then the Laplacian of f is

$$\Delta f = (*d*d - d*d*)f = *d*df = *d(i_{\nabla f}) = \mathrm{div}(\nabla f) = \nabla(\nabla f).$$

If $\mathscr{V}^1$ is a 1-form then

$$\begin{aligned} \Delta\mathscr{V} &= (*d*d - d*d*)\mathscr{V} = *d*d\mathscr{V} - d*d*\mathscr{V} = \langle\langle \nabla\times\nabla\times\mathbf{V}, \cdot \rangle\rangle - d(\nabla\cdot\mathbf{V}) \\ &= \langle\langle \nabla\times\nabla\times\mathbf{V} - \nabla(\nabla\cdot\mathbf{V}), \cdot \rangle\rangle = -\langle\langle \nabla^2\mathbf{V}, \cdot \rangle\rangle \end{aligned}$$

7.5.1 Cylindrical Polar coordinates

Consider cylindrical polar co-ordinates (r, θ, z). Then $x = r\cos\theta$, $y = r\sin\theta$ and hence $dx = \cos\theta dr - r\sin\theta d\theta$, $dy = \sin\theta dr + r\cos\theta d\theta$. From which we have

$$\begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} dr \\ d\theta \\ dz \end{bmatrix},$$

and

$$\begin{aligned} \mathrm{vol}^3 &= dx\wedge dy\wedge dz = r\,dr\wedge d\theta\wedge dz, \\ ds^2 &= dx\otimes dx + dy\otimes dy + dz\otimes dz = dr\otimes dr + r^2 d\theta\otimes d\theta + dz\otimes dz. \end{aligned}$$

We see that

$$\mathbf{e} = \begin{bmatrix} \partial_r & \frac{1}{r}\partial_\theta & \partial_z \end{bmatrix}$$

is an orthonormal basis and that

$$\begin{bmatrix}\partial_r & \frac{1}{r}\partial_\theta & \partial_z\end{bmatrix} = \begin{bmatrix}\partial_x & \partial_y & \partial_z\end{bmatrix}\begin{bmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{bmatrix}$$

Let

$$\begin{aligned}\mathscr{V} &= V_x dx + V_y dy + V_z dz = V_r dr + V_\theta d\theta + V_z dz,\\ \mathbf{V} &= \mathbf{V}_x\partial_x + \mathbf{V}_y\partial_y + \mathbf{V}_z\partial_z = \mathbf{V}_r\partial_r + \mathbf{V}_\theta\left(\frac{1}{r}\partial_\theta\right) + \mathbf{V}_z\partial_z.\end{aligned}$$

$$\begin{bmatrix}V_r & V_\theta & V_z\end{bmatrix} = \begin{bmatrix}V_x & V_y & V_z\end{bmatrix}\begin{bmatrix}\cos\theta & -r\sin\theta & 0\\ \sin\theta & r\cos\theta & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$\begin{bmatrix}\mathbf{V}_r\\ \mathbf{V}_\theta\\ \mathbf{V}_z\end{bmatrix} = \begin{bmatrix}\cos\theta & \sin\theta & 0\\ -\sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\mathbf{V}_x\\ \mathbf{V}_y\\ \mathbf{V}_z\end{bmatrix}$$

$$\begin{aligned}*\mathscr{V} &= i_{\mathbf{V}}\mathrm{vol}^3 = r\mathbf{V}_r d\theta\wedge dz + \mathbf{V}_\theta dz\wedge dr + r\mathbf{V}_z dr\wedge d\theta\\ \mathscr{V} &= \langle\langle \mathbf{V},\cdot\rangle\rangle = \mathbf{V}_r dr + r\mathbf{V}_\theta d\theta + \mathbf{V}_z dz,\end{aligned}$$

$$d\mathscr{V} = \left(\frac{\partial\mathbf{V}_z}{\partial\theta} - r\frac{\partial\mathbf{V}_\theta}{\partial z}\right)d\theta\wedge dz + \left(\frac{\partial\mathbf{V}_r}{\partial z} - \frac{\partial\mathbf{V}_z}{\partial r}\right)dz\wedge dr + \left(\frac{\partial(r\mathbf{V}_\theta)}{\partial r} - \frac{\partial\mathbf{V}_r}{\partial\theta}\right)dr\wedge d\theta.$$

Since

$$i_{\nabla\times\mathbf{V}}\mathrm{vol} = r(\nabla\times\mathbf{V})_r d\theta\wedge dz + (\nabla\times\mathbf{V})_\theta dz\wedge dr + r(\nabla\times\mathbf{V})_z dr\wedge d\theta$$

we have

$$\nabla\times\mathbf{V} = \left(\frac{1}{r}\frac{\partial\mathbf{V}_z}{\partial\theta} - \frac{\partial\mathbf{V}_\theta}{\partial z}\right)\partial_r + \left(\frac{\partial\mathbf{V}_r}{\partial z} - \frac{\partial\mathbf{V}_z}{\partial r}\right)\left(\frac{1}{r}\partial_\theta\right) + \frac{1}{r}\left(\frac{\partial(r\mathbf{V}_\theta)}{\partial r} - \frac{\partial\mathbf{V}_r}{\partial\theta}\right)\partial_z. \quad (7.58)$$

Hence

$$\begin{aligned}*d\mathscr{V} &= *i_{\nabla\times\mathbf{V}}\mathrm{vol} = \langle\langle\nabla\times\mathbf{V},\cdot\rangle\rangle\\ &= \left(\frac{1}{r}\frac{\partial\mathbf{V}_z}{\partial\theta} - \frac{\partial\mathbf{V}_\theta}{\partial z}\right)dr + r\left(\frac{\partial\mathbf{V}_r}{\partial z} - \frac{\partial\mathbf{V}_z}{\partial r}\right)d\theta + \frac{1}{r}\left(\frac{\partial(r\mathbf{V}_\theta)}{\partial r} - \frac{\partial\mathbf{V}_r}{\partial\theta}\right)dz\end{aligned}$$

Thus

$$\begin{aligned}d*d\mathscr{V} &= \left(\frac{1}{r}\left(\frac{\partial^2(r\mathbf{V}_\theta)}{\partial\theta\partial r} - \frac{\partial^2\mathbf{V}_r}{\partial\theta^2}\right) - r\left(\frac{\partial^2\mathbf{V}_r}{\partial z^2} - \frac{\partial^2\mathbf{V}_z}{\partial z\partial r}\right)\right)d\theta\wedge dz\\ &\quad + \left(\left(\frac{1}{r}\frac{\partial^2\mathbf{V}_z}{\partial z\partial\theta} - \frac{\partial^2\mathbf{V}_\theta}{\partial z^2}\right) - \left(\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial(r\mathbf{V}_\theta)}{\partial r}\right) - \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\mathbf{V}_r}{\partial\theta}\right)\right)\right)dz\wedge dr\end{aligned}$$

$$+\left(\left(\frac{\partial}{\partial r}\left(r\frac{\partial \mathbf{V}_r}{\partial z}\right)-\frac{\partial}{\partial r}\left(r\frac{\partial \mathbf{V}_z}{\partial r}\right)\right)-\left(\frac{1}{r}\frac{\partial^2 \mathbf{V}_z}{\partial \theta^2}-\frac{\partial^2 \mathbf{V}_\theta}{\partial \theta \partial z}\right)\right)dr\wedge d\theta$$

Since $i_{\nabla\times\nabla\times\mathbf{V}}\text{vol} = d*d*\mathscr{V}$ we have

$$\begin{aligned}\nabla\times\nabla\times\mathbf{V} &= \left(\frac{1}{r^2}\left(\frac{\partial^2(r\mathbf{V}_\theta)}{\partial\theta\partial r}-\frac{\partial^2\mathbf{V}_r}{\partial\theta^2}\right)-\left(\frac{\partial^2\mathbf{V}_r}{\partial z^2}-\frac{\partial^2\mathbf{V}_z}{\partial z\partial r}\right)\right)\partial_r\\ &\quad+\left(\left(\frac{1}{r}\frac{\partial^2\mathbf{V}_z}{\partial z\partial\theta}-\frac{\partial^2\mathbf{V}_\theta}{\partial z^2}\right)-\left(\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial(r\mathbf{V}_\theta)}{\partial r}\right)-\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\mathbf{V}_r}{\partial\theta}\right)\right)\right)\left(\frac{1}{r}\partial_\theta\right)\\ &\quad+\frac{1}{r}\left(\left(\frac{\partial}{\partial r}\left(r\frac{\partial\mathbf{V}_r}{\partial z}\right)-\frac{\partial}{\partial r}\left(r\frac{\partial\mathbf{V}_z}{\partial r}\right)\right)-\left(\frac{1}{r}\frac{\partial^2\mathbf{V}_z}{\partial\theta^2}-\frac{\partial^2\mathbf{V}_\theta}{\partial\theta\partial z}\right)\right)\partial_z \qquad (7.59)\end{aligned}$$

We also have

$$\begin{aligned}d*\mathscr{V} &= \left(\frac{1}{r}\frac{\partial(r\mathbf{V})}{\partial r}+\frac{1}{r}\frac{\partial\mathbf{V}_\theta}{\partial\theta}+\frac{\partial\mathbf{V}_z}{\partial z}\right)rdr\wedge d\theta\wedge dz,\\ *d*\mathscr{V} &= \left(\frac{1}{r}\frac{\partial(r\mathbf{V}_r)}{\partial r}+\frac{1}{r}\frac{\partial\mathbf{V}_\theta}{\partial\theta}+\frac{\partial\mathbf{V}_z}{\partial z}\right)=\nabla\cdot\mathbf{V},\\ d*d*\mathscr{V} &= \frac{\partial}{\partial r}(\nabla\cdot\mathbf{V})dr+\frac{\partial}{\partial\theta}(\nabla\cdot\mathbf{V})d\theta+\frac{\partial}{\partial z}(\nabla\cdot\mathbf{V})dz\\ *d*d*\mathscr{V} &= r\frac{\partial}{\partial r}(\nabla\cdot\mathbf{V})d\theta\wedge dz+\frac{1}{r}\frac{\partial}{\partial\theta}(\nabla\cdot\mathbf{V})dz\wedge dr+r\frac{\partial}{\partial z}(\nabla\cdot\mathbf{V})dr\wedge d\theta.\end{aligned}$$

Since $\nabla\cdot\mathbf{V} = *d*\mathscr{V}$ we have

$$\nabla\cdot\mathbf{V} = \left(\frac{1}{r}\frac{\partial(r\mathbf{V}_r)}{\partial r}+\frac{1}{r}\frac{\partial\mathbf{V}_\theta}{\partial\theta}+\frac{\partial\mathbf{V}_z}{\partial z}\right). \qquad (7.60)$$

Also since $d*d*\mathscr{V} = \langle\langle\nabla(\nabla\mathbf{V}),\cdot\rangle\rangle$ we have that

$$\nabla(\nabla\mathbf{V}) = \frac{\partial}{\partial r}(\nabla\cdot\mathbf{V})\partial_r+\frac{1}{r}\frac{\partial}{\partial\theta}(\nabla\cdot\mathbf{V})\left(\frac{1}{r}\partial_\theta\right)+\frac{\partial}{\partial z}(\nabla\cdot\mathbf{V})dz \qquad (7.61)$$

Let $\phi(r,\theta,z)$ be a function. Then

$$d\phi = \frac{\partial\phi}{\partial r}dr+\frac{\partial\phi}{\partial\theta}d\theta+\frac{\partial\phi}{\partial z}dz$$

Let $\nabla\phi$ be such that $\langle\langle\nabla\phi,\cdot\rangle\rangle = d\phi$. Hence

$$\nabla\phi = \frac{\partial\phi}{\partial r}\partial_r+\frac{1}{r}\frac{\partial\phi}{\partial\theta}\left(\frac{1}{r}\partial_\theta\right)+\frac{\partial\phi}{\partial z}\partial_z \qquad (7.62)$$

$$\begin{aligned}*d_s\phi &= i_{\nabla\phi}\text{vol} = r\frac{\partial\phi}{\partial r}d\theta\wedge dz+\frac{1}{r}\frac{\partial\phi}{\partial\theta}dz\wedge dr+r\frac{\partial\phi}{\partial z}dr\wedge d\theta,\\ d_s*d_s\phi &= \left(\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\phi}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2\phi}{\partial^2\theta}+\frac{\partial^2\phi}{\partial^2 z}\right)rdr\wedge d\theta\wedge dz,\\ \nabla^2\phi &= *d_s*d_s\phi = \left(\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\phi}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2\phi}{\partial^2\theta}+\frac{\partial^2\phi}{\partial^2 z}\right) \qquad (7.63)\end{aligned}$$

7.5.2 Spherical Polar coordinates

Consider spherical polar co-ordinates (r, φ, θ). Then $x = r\sin\varphi\cos\theta$, $y = r\sin\varphi\sin\theta$, $z = r\cos\varphi$ and hence

$$\begin{aligned}
dx &= \sin\varphi\cos\theta dr - r\sin\varphi\sin\theta d\theta + r\cos\varphi\cos\theta d\varphi \\
dy &= \sin\varphi\sin\theta dr + r\sin\varphi\cos\theta d\theta + r\cos\varphi\sin\theta d\varphi \\
dz &= \cos\varphi dr - r\sin\varphi d\varphi.
\end{aligned}$$

$$\begin{aligned}
\mathrm{vol}^3 &= dx\wedge dy\wedge dz = r^2\sin\varphi\, dr\wedge d\varphi\wedge d\theta, \\
ds^2 &= dx\otimes dx + dy\otimes dy + dz\otimes dz = dr\otimes dr + r^2 d\varphi\otimes d\varphi + r^2\sin^2\varphi d\theta\otimes d\theta.
\end{aligned}$$

Consider the orthonormal basis

$$\mathbf{e} = \begin{bmatrix} \partial_r & \frac{1}{r}\partial_\varphi & \frac{1}{r\sin\varphi}\partial_\theta \end{bmatrix}.$$

Let

$$\begin{aligned}
\mathscr{V} &= V_x dx + V_y dy + V_z dz = V_r dr + V_\varphi d\varphi + V_\theta d\theta, \\
\mathbf{V} &= \mathbf{V}_x\partial_x + \mathbf{V}_y\partial_y + \mathbf{V}_z\partial_z = \mathbf{V}_r\partial_r + \mathbf{V}_\varphi\left(\frac{1}{r}\partial_\varphi\right) + \mathbf{V}_\theta\left(\frac{1}{r\sin\varphi}\partial_\theta\right).
\end{aligned}$$

$$\begin{aligned}
*\mathscr{V} &= i_{\mathbf{V}}\mathrm{vol}^3 = r^2\sin\varphi\mathbf{V}_r d\varphi\wedge d\theta + r\sin\varphi\mathbf{V}_\varphi d\theta\wedge dr + r\mathbf{V}_\theta dr\wedge d\varphi \\
\mathscr{V} &= \langle\langle\mathbf{V},\cdot\rangle\rangle = \mathbf{V}_r dr + r\mathbf{V}_\varphi d\varphi + r\sin\varphi\mathbf{V}_\theta d\theta,
\end{aligned}$$

Hence

$$\begin{aligned}
d*\mathscr{V} &= \left(\frac{1}{r^2}\frac{\partial(r^2\mathbf{V}_r)}{\partial r} + \frac{1}{r\sin\varphi}\frac{\partial(\sin\varphi\mathbf{V}_\varphi)}{\partial\varphi} + \frac{1}{r\sin\varphi}\frac{\partial\mathbf{V}_\theta}{\partial\theta}\right) r^2\sin\varphi dr\wedge d\theta\wedge dz, \\
d\mathscr{V} &= \left(\frac{1}{r^2}\frac{\partial(r^2\mathbf{V}_r)}{\partial r} + \frac{1}{r\sin\varphi}\frac{\partial(\sin\varphi\mathbf{V}_\varphi)}{\partial\varphi} + \frac{1}{r\sin\varphi}\frac{\partial\mathbf{V}_\theta}{\partial\theta}\right) = \nabla\cdot\mathbf{V}, \\
d*d*\mathscr{V} &= \frac{\partial}{\partial r}(\nabla\cdot\mathbf{A})\,dr + \frac{\partial}{\partial\varphi}(\nabla\cdot\mathbf{V})\,d\varphi + \frac{\partial}{\partial\theta}(\nabla\cdot\mathbf{V})\,d\theta \\
d\mathscr{V} &= r\left(\frac{\partial(\sin\varphi\mathbf{V}_\theta)}{\partial\varphi} - \frac{\partial\mathbf{V}_\varphi}{\partial\theta}\right) d\varphi\wedge d\theta + \left(\frac{\partial\mathbf{V}_r}{\partial\theta} - \sin\varphi\frac{\partial(r\mathbf{V}_\theta)}{\partial r}\right) d\theta\wedge dr + \left(\frac{\partial(r\mathbf{V}_\varphi)}{\partial r} - \frac{\partial\mathbf{V}_r}{\partial\varphi}\right) dr\wedge d\varphi
\end{aligned}$$

Hence since $\mathrm{div}(\mathbf{V})\mathrm{vol} = d*\mathscr{V}$, and $i_{\nabla\times\mathbf{V}}\mathrm{vol} = d\mathscr{V}$ we have

$$\begin{aligned}
\mathrm{div}(\mathbf{V}) &= \nabla\cdot\mathbf{V} = \left(\frac{1}{r^2}\frac{\partial(r^2\mathbf{V}_r)}{\partial r} + \frac{1}{r\sin\varphi}\frac{\partial\mathbf{V}_\theta}{\partial\theta} + \frac{1}{r\sin\varphi}\frac{\partial(\sin\varphi\mathbf{V}_\varphi)}{\partial\varphi}\right), \\
\mathrm{curl}(\mathbf{V}) &= \nabla\times\mathbf{V} = \frac{1}{r\sin\varphi}\left(\frac{\partial(\sin\varphi\mathbf{V}_\theta)}{\partial\varphi} - \frac{\partial\mathbf{V}_\varphi}{\partial\theta}\right)\partial_r + \frac{1}{r}\left(\frac{1}{\sin\varphi}\frac{\partial\mathbf{V}_r}{\partial\theta} - \frac{\partial(r\mathbf{V}_\theta)}{\partial r}\right)\left(\frac{1}{r}\partial_\varphi\right) + \frac{1}{r}\left(\frac{\partial(r\mathbf{V}_\varphi)}{\partial r} - \frac{\partial\mathbf{V}_r}{\partial\varphi}\right)\left(\frac{1}{r\sin\varphi}\partial_\theta\right)
\end{aligned}$$

Hence

$$*d\mathscr{V} = \frac{1}{r\sin\varphi}\left(\frac{\partial(\sin\varphi\mathbf{V}_\theta)}{\partial\varphi} - \frac{\partial\mathbf{V}_\varphi}{\partial\theta}\right) dr + \left(\frac{1}{\sin\varphi}\frac{\partial\mathbf{V}_r}{\partial\theta} - \frac{\partial(r\mathbf{V}_\theta)}{\partial r}\right) d\varphi + \sin\varphi\left(\frac{\partial(r\mathbf{V}_\varphi)}{\partial r} - \frac{\partial\mathbf{V}_r}{\partial\varphi}\right) d\theta$$

Let $\phi(r,\varphi,\theta)$ be a function. Then

$$d\phi = \frac{\partial \phi}{\partial r}dr + \frac{\partial \phi}{\partial \varphi}d\varphi + \frac{\partial \phi}{\partial \theta}d\theta$$

Hence

$$\nabla\phi = \frac{\partial \phi}{\partial r}\partial_r + \frac{1}{r}\frac{\partial \phi}{\partial \varphi}\left(\frac{1}{r}\partial_\varphi\right) + \frac{1}{r\sin\varphi}\frac{\partial \phi}{\partial \theta}\left(\frac{1}{r\sin\varphi}\partial_\theta\right)$$

and since $*d\phi = i_{\nabla\phi}\mathrm{vol}$

$$*d\phi = r^2\sin\varphi\frac{\partial \phi}{\partial r}d\varphi\wedge d\theta + \sin\varphi\frac{\partial \phi}{\partial \varphi}d\theta\wedge dr + \frac{1}{\sin\varphi}\frac{\partial \phi}{\partial \theta}dr\wedge d\varphi$$

and

$$d*d\phi = \left(\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial \phi}{\partial r}\right) + \frac{1}{r^2\sin\varphi}\frac{\partial}{\partial \varphi}\left(\sin\varphi\frac{\partial \phi}{\partial \varphi}\right) + \frac{1}{r^2\sin^2\varphi}\frac{\partial^2 \phi}{\partial \theta^2}\right) r^2\sin\varphi dr\wedge d\varphi\wedge d\theta$$

Let

$$\mathbf{A} = \mathbf{A}_x\partial_x + \mathbf{A}_y\partial_y + \mathbf{A}_z\partial_z = \mathbf{A}_r\partial_r + \mathbf{A}_\varphi\left(\frac{1}{r}\partial_\varphi\right) + \mathbf{A}_\theta\left(\frac{1}{r\sin\varphi}\partial_\theta\right).$$

$$*\mathscr{A} = i_{\mathbf{A}}\mathrm{vol}^3 = r^2\sin\varphi\mathbf{A}_r d\varphi\wedge d\theta + r\sin\varphi\mathbf{A}_\varphi d\theta\wedge dr + r\mathbf{A}_\theta dr\wedge d\varphi$$

$$\mathscr{A} = \langle\langle \mathbf{A},\cdot\rangle\rangle = \mathbf{A}_r dr + r\mathbf{A}_\varphi d\varphi + r\sin\varphi\mathbf{A}_\theta d\theta,$$

Hence

$$d*\mathscr{A} = \left(\frac{1}{r^2}\frac{\partial(r^2\mathbf{A}_r)}{\partial r} + \frac{1}{r\sin\varphi}\frac{\partial(\sin\varphi\mathbf{A}_\varphi)}{\partial \varphi} + \frac{1}{r\sin\varphi}\frac{\partial \mathbf{A}_\theta}{\partial \theta}\right) r^2\sin\varphi dr\wedge d\theta\wedge dz,$$

$$*d*\mathscr{A} = \left(\frac{1}{r^2}\frac{\partial(r^2\mathbf{A}_r)}{\partial r} + \frac{1}{r\sin\varphi}\frac{\partial(\sin\varphi\mathbf{A}_\varphi)}{\partial \varphi} + \frac{1}{r\sin\varphi}\frac{\partial \mathbf{A}_\theta}{\partial \theta}\right) = \nabla\cdot\mathbf{A},$$

$$d*d*\mathscr{A} = \frac{\partial}{\partial r}(\nabla\cdot\mathbf{A})dr + \frac{\partial}{\partial \varphi}(\nabla\cdot\mathbf{A})d\varphi + \frac{\partial}{\partial \theta}(\nabla\cdot\mathbf{A})d\theta$$

$$*d*d*\mathscr{A} = r^2\sin\varphi\frac{\partial}{\partial r}(\nabla\cdot\mathbf{A})d\varphi\wedge d\theta + \sin\varphi\frac{\partial}{\partial \varphi}(\nabla\cdot\mathbf{A})d\theta\wedge dr + \frac{1}{\sin\varphi}\frac{\partial}{\partial \theta}(\nabla\cdot\mathbf{A})dr\wedge d\varphi$$

$$d\mathscr{A} = r\left(\frac{\partial(\sin\varphi\mathbf{A}_\theta)}{\partial \varphi} - \frac{\partial \mathbf{A}_\varphi}{\partial \theta}\right)d\varphi\wedge d\theta + \left(\frac{\partial \mathbf{A}_r}{\partial \theta} - \sin\varphi\frac{\partial(r\mathbf{A}_\theta)}{\partial r}\right)d\theta\wedge dr + \left(\frac{\partial(r\mathbf{A}_\varphi)}{\partial r} - \frac{\partial \mathbf{A}_r}{\partial \varphi}\right)dr\wedge d\varphi$$

Since $i_{\nabla\times\mathbf{A}}\mathrm{vol} = d\mathscr{A}$ we have

$$\nabla\times\mathbf{A} = \frac{1}{r\sin\varphi}\left(\frac{\partial(\sin\varphi\mathbf{A}_\theta)}{\partial \varphi} - \frac{\partial \mathbf{A}_\varphi}{\partial \theta}\right)\partial_r + \frac{1}{r}\left(\frac{1}{\sin\varphi}\frac{\partial \mathbf{A}_r}{\partial \theta} - \frac{\partial(r\mathbf{A}_\theta)}{\partial r}\right)\left(\frac{1}{r}\partial_\varphi\right) + \frac{1}{r}\left(\frac{\partial(r\mathbf{A}_\varphi)}{\partial r} - \frac{\partial \mathbf{A}_r}{\partial \varphi}\right)\left(\frac{1}{r\sin\varphi}\partial_\theta\right)$$

Since $*d\mathscr{A} = \langle\langle \nabla\times\mathbf{A},\cdot\rangle\rangle$ we have that

$$*d\mathscr{A}=\frac{1}{r\sin\varphi}\left(\frac{\partial(\sin\varphi\mathbf{A}_\theta)}{\partial\varphi}-\frac{\partial\mathbf{A}_\varphi}{\partial\theta}\right)dr+\left(\frac{1}{\sin\varphi}\frac{\partial\mathbf{A}_r}{\partial\theta}-\frac{\partial(r\mathbf{A}_\theta)}{\partial r}\right)d\varphi+\sin\varphi\left(\frac{\partial(r\mathbf{A}_\varphi)}{\partial r}-\frac{\partial\mathbf{A}_r}{\partial\varphi}\right)d\theta$$

$$\begin{aligned}d*d\mathscr{A}=&\left(\frac{\partial}{\partial\varphi}\left(\sin\varphi\left(\frac{\partial(r\mathbf{A}_\varphi)}{\partial r}-\frac{\partial\mathbf{A}_r}{\partial\varphi}\right)\right)-\left(\frac{1}{\sin\varphi}\frac{\partial^2\mathbf{A}_r}{\partial\theta^2}-\frac{\partial^2(r\mathbf{A}_\theta)}{\partial\theta\partial r}\right)\right)d\varphi\wedge d\theta\\&+\left(\frac{1}{r\sin\varphi}\left(\frac{\partial^2(\sin\varphi\mathbf{A}_\theta)}{\partial\theta\partial\varphi}-\frac{\partial^2\mathbf{A}_\varphi}{\partial\theta^2}\right)-\sin\varphi\left(\frac{\partial^2(r\mathbf{A}_\varphi)}{\partial r^2}-\frac{\partial^2\mathbf{A}_r}{\partial r\partial\varphi}\right)\right)d\theta\wedge dr\\&+\left(\left(\frac{1}{\sin\varphi}\frac{\partial^2\mathbf{A}_r}{\partial r\partial\theta}-\frac{\partial^2(r\mathbf{A}_\theta)}{\partial r^2}\right)-\frac{1}{r}\frac{\partial}{\partial\varphi}\left(\frac{1}{\sin\varphi}\left(\frac{\partial(\sin\varphi\mathbf{A}_\theta)}{\partial\varphi}-\frac{\partial\mathbf{A}_\varphi}{\partial\theta}\right)\right)\right)dr\wedge d\varphi\end{aligned}$$

7.6 Variational Calculus and Hamilton's Equations

Consider a curve $q:I\to M^n$ where $I=[t_1,t_2]$ and a 1-form Λ. We define the *action* of Λ over $q(t)$ as

$$S(q(t))=\int_{q(t)}\Lambda=\int_I q^*\Lambda=\int_I\Lambda(\dot q)dt.$$

Consider a variation of the curve $q:I\to M^n$ given by $q(\tau,\alpha)$ such that $q(\tau,0)=q(t)$. Then denote by $J(\tau,\alpha),T(\tau,\alpha)\in T_{q(\tau,\alpha)}M^n$ by

$$\begin{aligned}J(\tau,\alpha)&=q_*\left(\frac{\partial}{\partial\alpha}\right),\\T(\tau,\alpha)&=q_*\left(\frac{\partial}{\partial\tau}\right)\end{aligned}$$

For a fixed α denote the curve $q_\alpha(\tau)=q(\tau,\alpha)$. Consider the action of Λ along the curve $q_\alpha(\tau)=q(\tau,\alpha)$ given by

$$S(\alpha)=\int_{q_\alpha(\tau)}\Lambda=\int_I q_\alpha^*\Lambda=\int_I\Lambda(T(\tau,\alpha))d\tau.$$

$$\begin{aligned}\frac{dS}{d\alpha}&=\int_{q_\alpha(\tau)}\left(\frac{\partial\Lambda}{\partial\alpha}+\mathscr{L}_{J(\tau,\alpha)}\Lambda\right)=\int_{q_\alpha(\tau)}\left(\frac{\partial\Lambda}{\partial\alpha}+di_{J(\tau,\alpha)}\Lambda+i_{J(\tau,\alpha)}d\Lambda\right)\\&=\int_{\partial q_\alpha(\tau)}i_{J(\tau,\alpha)}\Lambda+\int_{q_\alpha(\tau)}\left(\frac{\partial\Lambda}{\partial\alpha}+i_{J(\tau,\alpha)}d\Lambda\right)\\&=(\Lambda(J(\tau_2,\alpha))-\Lambda(J(\tau_1,\alpha)))+\int_{q_\alpha(\tau)}\left(\frac{\partial\Lambda}{\partial\alpha}+i_{J(\tau,\alpha)}d\Lambda\right)\end{aligned}\tag{7.64}$$

Since Λ does not depend on the variation parameter α and we only consider variations that vanish at the end points. That is $J(\tau_1,0))=J(\tau_2,0)=0$. Then we have

$$\left.\frac{dS}{d\alpha}\right|_{\alpha=0}=\int_{q(t)}i_{J(\tau,0)}d\Lambda\tag{7.65}$$

If $S(q(t))$ is stationary at $q(t)$ then $\left.\frac{dS}{d\alpha}\right|_{\alpha=0} = 0$ for all variations $J(\tau,0)$ of $q(t)$ that vanish at the end points. Which implies that $i_J d\Lambda = 0$ for all J that are tangent to a variational surface that contains $q(t)$. Thus, since the tangent to the curve should also be on such a surface we should necessarily have $i_{\dot{q}} d\Lambda = 0$. In summary we have proved the following theorem:

Theorem 7.5. *The statement that $S(t)$ is stationary at $q(t)$ for all variations of $q(t)$ that vanish at the end points is equivalent to the statement that $d\Lambda \triangleq \Omega \equiv 0$ when restricted to $q(t)$. That is $i_{\dot{q}}\Omega = 0$.*

Consider $\mathbb{R} \times T^*M$ and the 1-form

$$\Lambda = p_i dq^i - H dt. \tag{7.66}$$

Consider a curve $c(t) = (t, q(t), p(t))$ and a variation $c(\tau,\alpha)$ such that $c(\tau,0) = c(t)$, $(t(\tau_1,\alpha), q(\tau_1,\alpha)) = (\tau_1, q(\tau_1))$ and $(t(\tau_2,\alpha), q(\tau_2,\alpha)) = (\tau_2, q(\tau_2))$. Let $J(\tau,\alpha) = \frac{\partial c}{\partial \alpha}$ and $T(\tau,\alpha) = \frac{\partial c}{\partial \tau}$.

Then

$$\begin{aligned}
\Omega = d\Lambda &= dp_i \wedge dq^i + dt \wedge dH \\
&= \frac{\partial H}{\partial q^i} dt \wedge dq^i + \left(\frac{\partial H}{\partial p_i} dt - dq^i\right) \wedge dp_i
\end{aligned} \tag{7.67}$$

From the above theorem since, $d\Lambda = 0$ when restricted to $(t, q(t), p(t))$, we have that $i_X \Omega = 0$ where $X = \frac{\partial}{\partial t} + \dot{q}^i \frac{\partial}{\partial q^i} + \dot{p}_i \frac{\partial}{\partial p_i}$ and hence that

$$\begin{aligned}
i_X \Omega &= -\left(\frac{\partial H}{\partial q^i} + \frac{\partial H}{\partial p_i}\right) dt + \left(\frac{\partial H}{\partial q^i} + \dot{p}_i\right) dq^i + \left(\frac{\partial H}{\partial p_i} - \dot{q}^i\right) dp_i \\
&= \left(\frac{\partial H}{\partial t} - \frac{dH}{dt}\right) dt + \left(\frac{\partial H}{\partial q^i} + \dot{p}_i\right) dq^i + \left(\frac{\partial H}{\partial p_i} - \dot{q}^i\right) dp_i \\
&= 0.
\end{aligned}$$

This shows that $i_X \Omega = 0$ is equivalent to the set of equations

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}, \tag{7.68}$$

$$\dot{q} = \frac{\partial H}{\partial p}, \tag{7.69}$$

$$\dot{p} = -\frac{\partial H}{\partial q}. \tag{7.70}$$

These are the *Hmilton's equations* for a system of interacting particles with Hamiltonian $H : \mathbb{R} \times T^*M \to \mathbb{R}$.

Theorem 7.6. *Let $c:[t_1,t_2]\to\mathbb{R}\times T^*M$ be a smooth curve that takes the coordinate form $c(t)=(t,q(t),p(t))$, $H:\mathbb{R}\times T^*M\to\mathbb{R}$ a smooth function,*

$$\Lambda = p_i dq^i - H dt, \tag{7.71}$$

a smooth 1-form and the action of the system along $c(t)$

$$S(c(t)) = \int_{c(t)} \Lambda. \tag{7.72}$$

The curve $c(t)$ is a stationery point of the action if and only if $c(t)$ satisfies the Hamilton's equations (7.68) – (7.70).

From (7.71) we also have that

$$\Omega \triangleq d\Lambda = dp_i \wedge dq^i - dH\wedge dt. \tag{7.73}$$

Since Ω by definition is exact it follows that $d\Omega=0$. The vector field

$$X_H = \frac{\partial}{\partial t} + \frac{\partial H}{\partial p} dq - \frac{\partial H}{\partial q} dp,$$

is defined to be the Hamiltonian vector field on $\mathbb{R}\times T^*M$. Since $i_{X_H}\Omega=0$ and $d\Omega=0$ it can be easily shown that

$$\mathscr{L}_{X_H}\Omega = 0. \tag{7.74}$$

Consider a particle with an associated Hamiltonian H in the absence of an electro-magnetic field. If in addition the particle has a charge e then it is also influenced by the electromagnetic potential $\mathscr{E}=\phi dt+\mathscr{A}_s$. Since by definition $d\mathscr{E}=\mathscr{F}$, and hence $d\mathscr{F}=0$ we are motivated to define a new action by augmenting Λ with $\mathscr{E}$ as follows.

$$\Lambda^e \triangleq \Lambda + e\mathscr{E} = (p_i + eA_i)dq^i - (H - e\phi)dt = p_i^e dq^i - H^e dt.$$

Then $\Omega^e \triangleq d\Lambda^e = \Omega + e\mathscr{F}$ where we recal that in units where the speed of light is equal to 1 $\mathscr{F} = \mathscr{E}_s\wedge dt + \mathscr{B}_s$ where $\mathscr{E}_s$ is the electrostatic 1-form and $\mathscr{B}_s$ is the magnetic 2-form associated with the fields $\mathbf{E}=(E_1,E_2,E_3)$ and $\mathbf{B}=(B_1,B_2,B_3)$. Here both $\mathscr{E}_s$ and $\mathscr{B}_s$ do not contain dt terms. Computing $i_X\Omega^e$ we find

$$\begin{aligned} i_X\Omega^e &= i_X\Omega + e i_X\mathscr{F} \\ &= i_X\Omega + e(i_{\dot q}\mathscr{E}_s)dt - e(\mathscr{E}_s - i_{\dot q}\mathscr{B}_s). \end{aligned}$$

Writing this out in Euclidean coordinates for space we have

$$i_X\Omega^e = \left(\frac{\partial H}{\partial t} - \frac{dH}{dt} + e\left(\dot q\cdot \mathbf{E}\right)\right)dt + \left(\frac{\partial H}{\partial q} + \dot p\right)dq + \left(\frac{\partial H}{\partial p} - \dot q\right)dp - e\left(\mathbf{E} + (\dot q\times\mathbf{B})\right)dq.$$

If the action $\mathscr{A}^e(c(t)) = \int_{c(t)} \Lambda^e$ is stationary along the curve $c(t) = (t, q(t), p(t))$ then we have seen that $c(t)$ should satisfy the equations

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + e\,(\dot{q} \cdot \mathbf{E}), \tag{7.75}$$

$$\dot{q} = \frac{\partial H}{\partial p}, \tag{7.76}$$

$$\dot{p} = -\frac{\partial H}{\partial q} + e\,(\mathbf{E} + (\dot{q} \times \mathbf{B})). \tag{7.77}$$

In fact experiments have verified that the behaviour of a classical particle influenced by the presence of an electromagnetic field is sufficiently accurately defined by these equations.

$$\begin{aligned}
i_X \Omega^e &= i_X \Omega + e i_X \mathscr{F} \\
&= i_X \Omega + e(i_{\dot{q}} \mathscr{E}_s) dt - e(\mathscr{E}_s - i_{\dot{q}} \mathscr{B}_s) \\
&= i_X \Omega + e\left(i_{\dot{q}} \mathbf{d}_s \phi - \frac{\partial i_{\dot{q}} \mathscr{A}_s}{\partial t} \right) dt - e\left(\left(\mathbf{d}_s \phi - \frac{\partial \mathscr{A}_s}{\partial t} \right) - i_{\dot{q}} \mathbf{d}_s \mathscr{A}_s \right) \\
&= i_X \Omega + e\left(\frac{d\phi}{dt} - \frac{\partial \phi}{\partial t} - \frac{\partial i_{\dot{q}} \mathscr{A}_s}{\partial t} \right) dt - e\left(\left(\mathbf{d}_s \phi - \frac{\partial \mathscr{A}_s}{\partial t} \right) - i_{\dot{q}} \mathbf{d}_s \mathscr{A}_s \right).
\end{aligned}$$

Define where $H^e \triangleq H(t, q, p^e - eA) - e\phi(t, q)$ and using the fact that $\dot{q} = \dfrac{\partial H}{\partial p}$ we see that

$$\begin{aligned}
\frac{\partial H^e}{\partial t} &= \frac{\partial}{\partial t}(H - e\phi) - e\frac{\partial H}{\partial p}\frac{\partial A}{\partial t} = \frac{\partial}{\partial t}(H - e\phi) - e\frac{\partial}{\partial t}(A \cdot \dot{q}), \\
\frac{\partial H^e}{\partial q} &= \frac{\partial}{\partial q}(H - e\phi) - e\frac{\partial A}{\partial q}\frac{\partial H}{\partial p} = \frac{\partial}{\partial q}(H - e\phi) - e\frac{\partial}{\partial q}(A \cdot \dot{q}).
\end{aligned}$$

Thus by direct computation we find that

$$\begin{aligned}
i_X \Omega^e &= \left(\frac{\partial}{\partial t}(H - e\phi - e\mathbf{A} \cdot \dot{q}) - \frac{d}{dt}(H - e\phi) \right) dt + \left(-e\left(\frac{\partial \phi}{\partial q} - \frac{\partial \mathbf{A}}{\partial t} + \dot{q} \times \nabla \mathbf{A} \right) + \frac{\partial H}{\partial q} + \dot{p} \right) dq + \left(\frac{\partial H}{\partial p} - \dot{q} \right) dp \\
&= \left(\frac{\partial}{\partial t}(H - e\phi - e\mathbf{A} \cdot \dot{q}) - \frac{d}{dt}(H - e\phi) \right) dt + \left(-e\left(-\frac{\partial \mathbf{A}}{\partial t} + \frac{d\mathbf{A}}{dt} + \dot{q} \times \nabla \mathbf{A} \right) + \frac{\partial}{\partial q}(H - e\phi) + \frac{d}{dt}(p + e\mathbf{A}) \right) dq + \left(\frac{\partial H}{\partial p} - \dot{q} \right) dp \\
&= \left(\frac{\partial}{\partial t}(H - e\phi - e\mathbf{A} \cdot \dot{q}) - \frac{d}{dt}(H - e\phi) \right) dt + \left(-e\left(-\frac{\partial \mathbf{A}}{\partial t} + \frac{d\mathbf{A}}{dt} + \dot{q} \times \nabla \mathbf{A} - \frac{\partial}{\partial q}(A \cdot \dot{q}) \right) + \frac{\partial}{\partial q}(H - e(A \cdot \dot{q}) - e\phi) + \frac{d}{dt}(p + e\mathbf{A}) \right) dq + \left(\frac{\partial H}{\partial p} - \dot{q} \right) dp \\
&= \left(\frac{\partial H^e}{\partial t} - \frac{dH^e}{dt} \right) dt + \left(\frac{\partial H^e}{\partial q} + \frac{dp^e}{dt} \right) dq + \left(\frac{\partial H^e}{\partial p^e} - \dot{q} \right) dp.
\end{aligned}$$

where we have used the fact

$$0 = \left(-\frac{\partial \mathbf{A}}{\partial t} + \frac{d\mathbf{A}}{dt} + \dot{q} \times \nabla \mathbf{A} - \frac{\partial}{\partial q}(\mathbf{A} \cdot \dot{q}) \right).$$

Thus we see that equations (7.75)–(7.77) are equivalent to

$$\frac{dH^e}{dt} = \frac{\partial H^e}{\partial t}, \tag{7.78}$$

$$\dot{q} = \frac{\partial H^e}{\partial p^e}, \tag{7.79}$$

$$\dot{p}^e = -\frac{\partial H^e}{\partial q}. \tag{7.80}$$

7.7 Lie Groups and Lie Algebras

A Lie group G is a group with a smooth manifold structure. That is, the group operation and the inversion map given below are both smooth [**?**, **?**, **?**].

$$\mu : G \times G \mapsto G \implies \mu(g,h) = gh, \tag{7.81}$$
$$i : G \mapsto G \implies i(g) = g^{-1}. \tag{7.82}$$

Denote by $\mathscr{G} \triangleq T_eG$ the tangent space to the identity element. It can be shown, using the product rule, that the derivatives of the above maps at the identity are given by

$$T_{(e,e)}\mu : \mathscr{G} \times \mathscr{G} \mapsto \mathscr{G} \implies T_{(e,e)}\mu(\zeta,\eta) = \zeta + \eta, \tag{7.83}$$
$$T_e i : \mathscr{G} \mapsto \mathscr{G} \implies T_e i(\zeta) = -\zeta. \tag{7.84}$$

Consider the conjugation map $\mathrm{I} : G \times G \to G$ given by $\mathrm{I}_g(h) \triangleq ghg^{-1}$. It is easy to show that I_g is a group homomorphism for every $g \in G$. Its derivative map at the identity $T_e\mathrm{I}_g : \mathscr{G} \mapsto \mathscr{G}$ is called the *adjoint map*

$$\mathrm{Ad}_g\zeta \triangleq T_e\mathrm{I}_g(\zeta) = \frac{d}{dt}\,|_{t=0}\, gh(t)g^{-1}, \tag{7.85}$$

where $h(t)$ is a smooth curve such that $h(0) = e$ and $\dot{h}(0) = \zeta$. Notice that $\mathrm{Ad}_g \in GL(\mathscr{G})$ and that the map $\mathrm{Ad} : G \mapsto GL(\mathscr{G})$ is a group homomorphism called the *adjoint representation*. The derivative of this map at the identity is then a linear map from $\mathscr{G}$ to the linear maps from $\mathscr{G}$ to $\mathscr{G}$ that is denoted by $L(\mathscr{G},\mathscr{G})$. This map is given by

$$\mathrm{ad}_\zeta\eta \triangleq \frac{d}{ds}\,|_{s=0}\, \mathrm{Ad}_{g(s)}\eta, \tag{7.86}$$

where $g(s)$ is a smooth curve such that $g(0) = e$ and $\dot{g}(0) = \zeta$. It is easy to verify that this $\mathrm{ad} : \mathscr{G} \times \mathscr{G} \mapsto \mathscr{G}$ operator is bi-linear.

Consider the conjugation map $C : G \times G \to G$ that is given by $C_g(h) = ghg^{-1}h^{-1}$. Taking the derivative of this map with respect to h at the identity we have

$$T_eC_g(\eta) = \frac{d}{dt}\,|_{t=0}\, C_g(h(t)) = (\mathrm{Ad}_g - I)\eta, \tag{7.87}$$

and taking the derivative of the conjugation map with respect to g at the identity we have

$$(T_eC(h))(\zeta) = \frac{d}{ds}\,|_{s=0}\, C_{g(s)}(h) = (I - \mathrm{Ad}_h)\zeta, \tag{7.88}$$

Differentiating the above maps twice we have

$$\frac{d}{ds}\,|_{s=0}\, \frac{d}{dt}\,|_{t=0}\, C_{g(s)}(h(t)) = \mathrm{ad}_\zeta\eta,$$
$$\frac{d}{dt}\,|_{t=0}\, \frac{d}{ds}\,|_{s=0}\, C_{g(s)}(h(t)) = -\mathrm{ad}_\eta\zeta,$$

and hence that $\mathrm{ad}_\zeta \eta$ gives a skew symmetric product structure on $\mathscr{G}$ that generalizes the cross-product in $\mathbb{R}^3$. We will denote this product structure on $\mathscr{G}$ by

$$[\zeta,\eta] \triangleq \mathrm{ad}_\zeta \eta. \tag{7.89}$$

What we showed above is that the product structure $[\cdot,\cdot]$ is anti-commutative or anti-symmetric. That is

$$[\zeta,\eta] = -[\eta,\zeta]. \tag{7.90}$$

Consider a group homomorphism $\phi : G \mapsto G'$. Since $\phi \cdot \mathrm{I}_g(h) = \mathrm{I}_{\phi(g)}\phi(h)$ or explicitly $\phi(ghg^{-1}) = \phi(g)\phi(h)\phi(g^{-1})$ we see that

$$T_e\phi \cdot \mathrm{Ad}_g\, \eta = \mathrm{Ad}_{\phi(g)} \cdot T_e\phi\, \eta. \tag{7.91}$$

Note that the Ad_g on the left is on G while the one on the right is on G'. Differentiating it again at the identity we have

$$T_e\phi \cdot \mathrm{ad}_\zeta\, \eta = \mathrm{ad}_{T_e\phi\, \zeta} \cdot T_e\phi\, \eta. \tag{7.92}$$

This is equivalent to

$$T_e\phi \cdot [\zeta,\eta]_{\mathscr{G}} = [T_e\phi\, \zeta, T_e\phi\, \eta]_{\mathscr{G}'}. \tag{7.93}$$

Consider the Lie group $GL(\mathscr{G})$ where $T_I GL = L(\mathscr{G},\mathscr{G})$ is the Lie algebra of $GL(\mathscr{G})$ with $[\zeta,\eta] = \mathrm{ad}_\zeta \eta = \zeta\eta - \eta\zeta$ where now $\zeta,\eta \in L(\mathscr{G},\mathscr{G})$ are matrices. Note that $\mathrm{Ad} : G \mapsto GL(\mathscr{G})$ is a group homomorphism and hence that $\mathrm{Ad}_g \in GL(\mathscr{G})$ is a matrix. Then (7.91) becomes

$$\mathrm{ad}_{\mathrm{Ad}_g\, \eta} = \mathrm{Ad}_g \mathrm{ad}_\eta \mathrm{Ad}_{g^{-1}}.$$

Letting this operate on $\mathrm{Ad}_g\, \eta$ we have

$$[\mathrm{Ad}_g\, \eta, \mathrm{Ad}_g\, \zeta] = \mathrm{Ad}_g[\eta,\zeta]. \tag{7.94}$$

Furthermore $T_e\mathrm{Ad}_g(\zeta) = \mathrm{ad}_\zeta = [\zeta,] \in GL(\mathscr{G})$ and then (7.92) becomes

$$\mathrm{ad}_{[\zeta,\eta]} = [\mathrm{ad}_\zeta, \mathrm{ad}_\eta]_{\mathscr{G}'},$$

and when operating on ν we have

$$\begin{aligned}[][[\zeta,\eta],\nu]_{\mathscr{G}} &= [\mathrm{ad}_\zeta, \mathrm{ad}_\eta]_{\mathscr{G}'}\nu = \mathrm{ad}_\zeta \mathrm{ad}_\eta \nu - \mathrm{ad}_\eta \mathrm{ad}_\zeta \nu,\\ &= [\zeta,[\eta,\nu]] - [\eta,[\zeta,\nu]].\end{aligned}$$

Thus from the anti-commutativity of the product $[\cdot,\cdot]$ we see that the product also satisfies the Jacobi rule

$$0 = [\zeta,[\eta,\nu]] + [\eta,[\nu,\zeta]] + [\nu,[\zeta,\eta]]. \tag{7.95}$$

A vector space $\mathscr{V}$ with a product $[\cdot,\cdot]:\mathscr{V}\times\mathscr{V}\mapsto\mathscr{V}$ that satisfies the anti-comutavity (7.90) and the Jacobi rule (7.95) is called a Lie-algbra. Thus $\mathscr{G}=T_eG$ equipped with the product $[\zeta,\eta]=\mathrm{ad}_\zeta\eta$ is referred to as the Lie algebra of the Lie group G.

For every $g\in G$ we can define the left multiplication and right multiplication maps $L_g:G\mapsto G$ and $R_g:G\mapsto G$ that are respectively given by $L_g(h)=gh$ and $R_g(h)=hg$. Since $L_{g_1g_2}=L_{g_1}\cdot L_{g_2}$ and $L_{g_1g_2}=R_{g_2}\cdot R_{g_1}$ we see that left multiplication is a homomorphism while the right multiplication is an anti-homomorphism.

A vectorfield X is said to be left-invariant if it satisfies $X(gh)=T_hL_g(X(h))$ for all $g,h\in G$ while it is said to be right invariant if $X(hg)=T_hR_g(X(h))$ for all $g,h\in G$. Notice that this means that both left-invariant and right-invariant vectorfields are completely determined by their value at the identity. We will use the notations $X^L_\zeta(g)\triangleq T_eL_g(\zeta)$ and $X^R_\zeta(g)\triangleq T_eR_g(\zeta)$ for $\zeta\in\mathscr{G}$ denote the left and right invariant vector fields generated by respectively left and right translating $\zeta\in\mathscr{G}=T_eG$ to T_gG.

Let $\Phi^t_{X^L_\zeta}:G\mapsto G$ be the flow of $X^L_\zeta(g)$ defined for $t\in(-\varepsilon,\varepsilon)$ for some $\varepsilon\in\mathbb{R}$. From the existence of solutions such an ε is guaranteed to exist. By definition $\Phi^t_{X^L_\zeta}(g)$ is the solution of $X^L_\zeta(g)$ through g. Thus we have $\frac{d}{dt}\,|_{t=0}\,\Phi^t_{X^L_\zeta}(g)=X^L_\zeta(g)$ and $\frac{d}{dt}\,|_{t=0}\,\Phi^t_{X^L_\zeta}(e)=\zeta$ and

$$\frac{d}{dt}\,|_{t=0}\,L_g(\Phi^t_{X^L_\zeta}(e))=T_eL_g(\zeta)=X^L_\zeta(g)$$

then the uniqueness of solutions imply that

$$\Phi^t_{X^L_\zeta}(g)=L_g(\Phi^t_{X^L_\zeta}(e))=R_{\Phi^t_{X^L_\zeta}(e)}g.$$

Similarly we can show that flow $\Phi^t_{X^R_\zeta}:G\mapsto G$ of $X^R_\zeta(g)$ satisfies

$$\Phi^t_{X^R_\zeta}(g)=R_g(\Phi^t_{X^R_\zeta}(e))=L_{\Phi^t_{X^R_\zeta}(e)}g.$$

This implies that, for $s\in(-\varepsilon,\varepsilon)$, $L_{\Phi^t_{X^L_\zeta}(e)}(\Phi^s_{X^L_\zeta}(e))=\Phi^s_{X^L_\zeta}(\Phi^t_{X^L_\zeta}(e))=\Phi^{s+t}_{X^L_\zeta}(e)$ and hence the map $\Phi^{s+t}_{X^L_\zeta}(e)$ is defined and therefore we see that the flow is in fact defined for all $t\in(-\infty,\infty)$. It also implies that the map

$$\Phi(e):\mathbb{R}\mapsto G\quad\Longrightarrow\quad t\mapsto\Phi^t_{X^L_\zeta}(e)=g(t)\in G \tag{7.96}$$

is a 1-parameter group homomorphism and that when viewed as a curve on G satisfies the property that its derivative at e is equal to ζ. Similarly one can show that $\Phi^t_{X^R_\zeta}(e)$ is a 1-parameter group homomorphism such that when viewed as a curve on G satisfies the property that its derivative at e is equal to ζ.

Conversely assume that $g:\mathbb{R}\mapsto G$ is a group homomorphism such that its derivative at the identity is equal to ζ. From the homomorphism property it follows that

$$g(t+s) = g(t)g(s) = g(s)g(t)$$

Since

$$\begin{aligned}\frac{d}{dt}g(t) &= \frac{d}{ds}\,|_{s=0}\, g(t+s) = \frac{d}{ds}\,|_{s=0}\, (g(t)g(s)) = \frac{d}{ds}\,|_{s=0}\, (g(s)g(t)) \\ &= T_eL_{g(t)}\zeta = X^L_\zeta(g(t)) \\ &= T_eR_{g(t)}\zeta = X^R_\zeta(g(t)).\end{aligned}$$

Thus if $t \mapsto g(t)$ is a homomorphism such that its derivative at e is ζ then it is the unique solution curve of both $X^L_\zeta(g(t))$ and $X^R_\zeta(g(t))$ passing through e. Thus we have proven the following:

Theorem 7.7. *The group homomorphism $g : \mathbb{R} \mapsto G$ such that $\frac{d}{dt}g(t)\,|_{t=0} = \zeta$ is unique and coincides with the solution curve $\Phi^t_{X^L_\zeta}(e)$ of $X^L_\zeta(g)$ and $\Phi^t_{X^R_\zeta}(e)$ of $X^R_\zeta(g)$ passing through the identity element e.*

The curve $g(t)$ that results in this fashion will be denoted as follows:

$$\exp(\zeta t) \triangleq \Phi^t_\zeta(e) \triangleq \Phi^t_{X^L_\zeta}(e) = \Phi^t_{X^R_\zeta}(e). \tag{7.97}$$

Lets consider the derivative of the flows $\Phi^t_{X^L_\zeta} : G \mapsto G$ and $\Phi^t_{X^R_\zeta} : G \mapsto G$ at the point $g \in G$.

$$T_g\Phi^t_{X^L_\zeta} \cdot T_eL_g\eta \triangleq \frac{d}{ds}\,|_{s=0}\, \Phi^t_{X^L_\zeta}(g\exp(\eta s)) = \frac{d}{ds}\,|_{s=0}\, R_{\Phi^t_{X^L_\zeta}(e)}g\exp(\eta s) = T_gR_{\Phi^t_{X^L_\zeta}(e)} \cdot T_eL_g\eta$$

$$T_g\Phi^t_{X^R_\zeta} \cdot T_eR_g\eta \triangleq \frac{d}{ds}\,|_{s=0}\, \Phi^t_{X^R_\zeta}(\exp(\eta s)\,g) = \frac{d}{ds}\,|_{s=0}\, L_{\Phi^t_{X^R_\zeta}(e)}\exp(\eta s)\,g = T_gL_{\Phi^t_{X^R_\zeta}(e)} \cdot T_eR_g\eta$$

From these two properties we see that

$$\begin{aligned}[X^L_\zeta, X^L_\eta]_{\mathscr{X}}(e) &\triangleq \lim_{t\to 0}\frac{T_{\Phi^t_{X^L_\zeta}(e)}\Phi^{-t}_{X^L_\zeta} \cdot X^L_\eta(\Phi^t_{X^L_\zeta}(e)) - \eta}{t} = \lim_{t\to 0}\frac{T_{\Phi^t_{X^L_\zeta}(e)}\Phi^{-t}_{X^L_\zeta} \cdot T_eL_{\Phi^t_{X^L_\zeta}(e)}\eta - \eta}{t} \\ &= \lim_{t\to 0}\frac{T_{\Phi^t_{X^L_\zeta}(e)}R_{\Phi^{-t}_{X^L_\zeta}(e)} \cdot T_eL_{\Phi^t_{X^L_\zeta}(e)}\eta - \eta}{t} \\ &= \frac{d}{dt}\,|_{t=0}\, (\mathrm{Ad}_{\Phi^t_{X^L_\zeta}(e)}\eta - \eta) \\ &= \mathrm{ad}_\zeta\eta = [\zeta, \eta]_{\mathscr{G}}.\end{aligned}$$

Similarly we see that

$$
\begin{aligned}
[X^R_\zeta, X^R_\eta]_{\mathscr{X}}(e) &\triangleq \lim_{t\to 0} \frac{T_{\Phi^t_{X^R_\zeta}(e)}\Phi^{-t}_{X^R_\zeta}\cdot X^R_\eta(\Phi^t_{X^R_\zeta}(e)) - \eta}{t} = \lim_{t\to 0}\frac{T_{\Phi^t_{X^R_\zeta}(e)}\Phi^{-t}_{X^R_\zeta}\cdot T_eR_{\Phi^t_{X^R_\zeta}(e)}\eta - \eta}{t}\\
&= \lim_{t\to 0}\frac{T_{\Phi^t_{X^R_\zeta}(e)}L_{\Phi^{-t}_{X^R_\zeta}(e)}\cdot T_eR_{\Phi^t_{X^R_\zeta}(e)}\eta - \eta}{t}\\
&= \frac{d}{dt}\,|_{t=0}\left(\mathrm{Ad}_{\Phi^{-t}_{X^R_\zeta}(e)}\eta - \eta\right)\\
&= -\mathrm{ad}_\zeta\eta = -[\zeta,\eta]_{\mathscr{G}}.
\end{aligned}
$$

We can also show similarly that

$$[X^L_\zeta, X^R_\zeta]_{\mathscr{X}} = 0.$$

Since $L_{\Phi^t_\zeta(e)}(\Phi^{-t}_\zeta(e)) = e$ for all $t\in\mathbb{R}$, we have that

$$
\begin{aligned}
i(\Phi^t_\zeta(e)) &= \Phi^{-t}_\zeta(e) = \exp(-\zeta t)\\
\frac{d}{dt}\,|_{t=0}\,\Phi^{-t}_\zeta(e) &= -\zeta\\
T_gi\cdot X^L_\zeta(g) &= \frac{d}{dt}\,|_{t=0}\, i(L_g\exp\zeta t) = \frac{d}{dt}\,|_{t=0}\left(R_{i(g)}\cdot\exp(-\zeta t)\right) = -X^R_\zeta(i(g)).
\end{aligned}
$$

The last equality means that

$$
\begin{aligned}
X^L_\zeta(i(g)) &= -T_gi\cdot X^R_\zeta(g).\\
X^L_\zeta(g) &= -T_{i(g)}i\cdot X^R_\zeta(i(g)).
\end{aligned}
$$

We also see that

$$X^L_\zeta(g) = T_eL_g\cdot\zeta = T_eL_g\cdot T_gR_{i(g)}\cdot X^R_\zeta(g) = T_g(L_g\cdot R_{i(g)})\cdot\zeta^R(g) = \mathrm{Ad}_gX^R_\zeta(g)$$

Consider the Lie algebra of vectorfields $X(g)$ equipped with the Jacobi-Lie bracket of vectorfields. We will denote this Lie algebra by $(\mathscr{X},[\cdot,\cdot]_{\mathscr{X}})$. Consider the subspaces of $\mathscr{X}$ that corresponds to the set of left invariant vectorfields that we will denote by $\mathscr{X}^L$ and the set of right invariant vectorfields that we will denote by $\mathscr{X}^R$. Since for all $g\in G$ the left translation and right translation maps respectively denoted by, $L_g : G\mapsto G$ and $R_g : G\mapsto G$ are isomomorphisms on the manifold G we have

$$
\begin{aligned}
[X^L_\zeta(gh), X^L_\eta(gh)]_{\mathscr{X}} &= [T_hL_g\cdot X^L_\zeta(h), T_hL_g\cdot X^L_\zeta(h)]_{\mathscr{X}} = T_hL_g\cdot[X^L_\zeta(h), X^L_\eta(h)]_{\mathscr{X}}\\
[X^R_\zeta(hg), X^R_\eta(hg)]_{\mathscr{X}} &= [T_hR_g\cdot X^R_\zeta(h), T_hR_g\cdot X^R_\zeta(h)]_{\mathscr{X}} = T_hR_g\cdot[X^R_\zeta(h), X^R_\eta(h)]_{\mathscr{X}}
\end{aligned}
$$

Which says that the Jacobi Lie bracket of left and right invariant vectorfields are also respectively left and right invariant. That is

$$[X^L_\zeta, X^L_\eta]_{\mathscr{X}} = T_eL_g\cdot[X^L_\zeta, X^L_\eta]_{\mathscr{X}} = T_eL_g\cdot[\zeta,\eta]_{\mathscr{G}} = X^L_{[\zeta,\eta]_{\mathscr{G}}}$$

$$[X^R_\zeta, X^R_\eta]_{\mathscr{X}} = T_e R_g \cdot [X^R_\zeta, X^R_\eta]_{\mathscr{X}} = -T_e R_g \cdot [\zeta, \eta]_{\mathscr{G}} = -X^R_{[\zeta,\eta]_{\mathscr{G}}}$$

Which implies that $\mathscr{X}^L$ and $\mathscr{X}^R$ are subalgebras of $(\mathscr{X}, [\cdot,\cdot]_{\mathscr{X}})$ and that the vectorspace isomorphism defined by

$$^L : \mathscr{G} \mapsto \mathscr{X}^L \quad \Longrightarrow \zeta \mapsto X^L_\zeta$$

is a Lie algebra homomorphism while the vectorspace isomorphism defined by

$$^R : \mathscr{G} \mapsto \mathscr{X}^R \quad \Longrightarrow \zeta \mapsto X^R_\zeta$$

is a Lie algebra anti-homomorphism.

Consider the group homomorphism $\phi : G \mapsto H$ and the 1-parameter subgroup $h(t) = \phi(\exp(\zeta t))$ on H. Then since

$$\frac{d}{dt}|_{t=0} h(t) = \frac{d}{dt}|_{t=0} \phi(\exp(\zeta t)) = T_e\phi \cdot \zeta.$$

we see that $h(t) = \phi(\exp(\zeta t))$ is the unique 1-parameter subgroup on H such that $\dot{h}(0) = T_e\phi \cdot \zeta$. Thus we have that

$$\phi(\exp(t\zeta)) = \exp(tT_e\phi \cdot \zeta). \tag{7.98}$$

In the special case where $\phi = \mathrm{Ad}$, we have that

$$h(t) = \mathrm{Ad}_{\exp(\zeta t)} = \exp(t\,\mathrm{ad}_\zeta) = I + t\,\mathrm{ad}_\zeta + \frac{t^2}{2!}\mathrm{ad}^2_\zeta + \cdots \tag{7.99}$$

In the special case where $\phi : GL(n) \mapsto \mathbb{R}\backslash\{0\}$ given by $\phi(A) = \det(A)$, we have that

$$h(t) = \det(\exp(At)) = \exp(tT_I \det A) = \exp(\mathrm{trace}(tA)) \tag{7.100}$$

Let

$$\Gamma(s,t) \triangleq \exp(-s\zeta(t)) \frac{\partial}{\partial t} \exp(s\zeta(t))$$

Then

$$\begin{aligned}\frac{\partial \Gamma}{\partial s} &= \exp(-s\zeta(t))(-\zeta(t)\frac{\partial}{\partial t}\exp(s\zeta(t)) + \exp(-s\zeta(t))\frac{\partial}{\partial t}(\zeta(t)\exp(s\zeta(t))) \\ &= \exp(-s\zeta(t))\frac{d\zeta}{dt}(\exp(s\zeta(t))) = \mathrm{Ad}_{\exp(-s\zeta(t))}\frac{d\zeta}{dt} = \exp\left(-s\,\mathrm{ad}_{\zeta(t)}\right)\frac{d\zeta}{dt}\end{aligned}$$

Thus

$$\begin{aligned} e^{-\zeta(t)}\frac{\partial}{\partial t}e^{\zeta(t)} = \Gamma(1,t) &= \int_0^1 \frac{\partial \Gamma}{\partial s}\,ds = \left(\int_0^1 \exp\left(-s\,\mathrm{ad}_{\zeta(t)}\right) ds\right)\frac{d\zeta}{dt} \\ &\triangleq \left(\frac{I - e^{-\mathrm{ad}_\zeta}}{\mathrm{ad}_\zeta}\right)\frac{d\zeta}{dt}\end{aligned}$$

From which we get

$$\frac{d}{dt}e^{\zeta(t)} = e^{\zeta(t)}\left(\frac{I-e^{-\mathrm{ad}_\zeta}}{\mathrm{ad}_\zeta}\right)\frac{d\zeta}{dt} \tag{7.101}$$

Let $e^{\chi(t)} = e^{\zeta}e^{t\eta}$. Then

$$\eta = e^{-\chi(t)}\frac{d}{dt}e^{\chi(t)} = \left(\frac{I-e^{-\mathrm{ad}_\chi}}{\mathrm{ad}_\chi}\right)\frac{d\chi(t)}{dt}.$$

Thus formally

$$\frac{d\chi(t)}{dt} = \left(\frac{\mathrm{ad}_\chi}{I-e^{-\mathrm{ad}_\chi}}\right)\eta.$$

Hence integrating from 0 to 1 with respect to t we have

$$\log\left(e^{\zeta}e^{\eta}\right) = \chi(1) = \zeta + \left(\int_0^1 \frac{\mathrm{ad}_\chi}{I-e^{-\mathrm{ad}_\chi}}\,dt\right)\eta.$$

The formal series in the integral is defined as

$$\frac{\mathrm{ad}_\chi}{I-e^{-\mathrm{ad}_\chi}} = \psi(e^{\mathrm{ad}_\chi}) = \psi(\mathrm{Ad}_{e^\chi}) = \psi(\mathrm{Ad}_{e^\zeta e^{t\eta}}) = \psi(e^{\mathrm{ad}_\zeta}e^{t\,\mathrm{ad}_\eta})$$

where

$$\psi(w) \triangleq \frac{w\log w}{w-1} = 1 + \sum_{m=1}^{\infty}\frac{(-1)^{m+1}}{m(m+1)}(w-1)^m, \quad ||w|| < 1.$$

Thus we have the Baker-Campbell-Hausdorff formula

$$\log\left(e^{X}e^{Y}\right) = X + \left(\int_0^1 \psi(e^{\mathrm{ad}_X}e^{t\,\mathrm{ad}_Y})\,dt\right)Y \tag{7.102}$$

$$= X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}\left([X,[X,Y]] + [Y,[Y,X]]\right) \tag{7.103}$$

$$-\frac{1}{24}[Y,[X,[X,Y]]] \tag{7.104}$$

$$-\frac{1}{720}\left([Y,[Y,[Y,[Y,X]]]] + [X,[X,[X,[X,Y]]]]\right) \tag{7.105}$$

$$+\frac{1}{360}\left([X,[Y,[Y,[Y,X]]]] + [Y,[X,[X,[X,Y]]]]\right) \tag{7.106}$$

$$+\frac{1}{120}\left([Y,[X,[Y,[X,Y]]]] + [X,[Y,[X,[Y,X]]]]\right) + \cdots. \tag{7.107}$$

A smooth vector field $X^L \in \mathfrak{X}(G)$ on a Lie-group G is called left-invariant if,

$$X^L(L_g h) = (L_g)_* X^L(h) \triangleq g \cdot X^L(h)$$

for every $g,h \in G$ where $(L_g)_* : T_hG \mapsto T_{gh}G$ is the derivative map of $L_g : G \mapsto G$. Similarly a smooth vector field $X^R \in \mathfrak{X}(G)$ on a Lie-group G is called right-invariant if,

$$X^R(R_g h) = (R_g)_* X^R(h) \triangleq X^R(h) \cdot g.$$

Let $\mathfrak{X}_L(G), \mathfrak{X}_R(G) \subset \mathfrak{X}(G)$ be the space of left and right-invariant vector fields on G respectively. Show that $\mathfrak{X}_R(G)$ is isomorphic to $\mathfrak{X}_L(G)$, i.e. the space of left-invariant vector fields on G.............

Let $\mathfrak{X}(G)$ be the space of vector fields on Lie-group G. Let $\mathfrak{X}_L(G), \mathfrak{X}_R(G) \subset \mathfrak{X}(G)$ be the spaces of left invariant and right invariant vector fields on G respectively. Then for a smooth vector field $X^L \in \mathfrak{X}_L(G)$ we have $\forall h, g \in G$,

$$X^L(L_g h) = (L_g)_* X^L(h) = g \cdot X^L(h).$$

In particular

$$X^L(h) = (L_h)_* X^L(e) = h \cdot X^L(e). \tag{7.108}$$

Define the vector field

$$X^R(h) \triangleq (R_h)_* X^L(e) \triangleq X^L(e) \cdot h, \tag{7.109}$$

for all $h \in G$. Then

$$X^R(R_g h) = X^R(hg) = \left(R_{hg}\right)_* X^L(e) = (R_g)_* (R_h)_* X^L(e) = (R_g)_* X^R(h),$$

and hence that $X^R(h)$ defined by (7.109) is such that

$$X^R(h) \in \mathfrak{X}_R(G).$$

That is we have shown that the vector field $X^R(h)$ defined by (7.109) is a right invariant vector field.

Furthermore from (7.108) and (7.109) we also have that

$$X^R(h) \triangleq (R_h)_* X^L(e) = (R_h)_* (L_{h^{-1}})_* X^L(h).$$

Notice that $(R_h)_* (L_{h^{-1}})_*$ is the inverse of the derivative of the map $I_h : G \mapsto G$ defined such that $I_h(g) \triangleq L_h R_{h^{-1}} g \triangleq R_{h^{-1}} L_h g \triangleq hgh^{-1}$. That is $(R_h)_* (L_{h^{-1}})_* = (I_{h^{-1}})_* = ((I_h)_*)^{-1}$.

Then we have that,

$$X^R(h) \triangleq (R_h)_* X^L(e) = (R_h)_* (L_{h^{-1}})_* X^L(h) = (I_{h^{-1}})_* X^L(h) = h^{-1} \cdot X^L(h) \cdot h \in \mathfrak{X}_R(G).$$

Thus we have shown

$$(I_{h^{-1}})_* : \mathfrak{X}_L(G) \mapsto \mathfrak{X}_R(G).$$

Since $((I_h)_*)^{-1} = (I_{h^{-1}})_*$ the map $(I_{h^{-1}})_*$ is an isomorphism this proves that $\mathfrak{X}_R(G)$ is isomorphic to $\mathfrak{X}_L(G)$.

Let

$$\dot{g} = g \cdot u$$

and $\phi : G \mapsto GL(\mathscr{G})$ be a group homomorphism and $f : GL(\mathscr{G}) \mapsto \mathbb{R}$ a polar Morse function on $GL(\mathscr{G})$ with a unique minimum at $I_{n\times n}$. This induces a function $f_\phi : G \mapsto \mathbb{R}$ defined by $f_\phi(g) = f(\phi(g))$.

7.8 Group Actions

Let $\phi : G \times P \to P$, denoted by $\phi(g,p) \triangleq g \cdot p$ be a left or right action of G on P. Let $\mathscr{X}_P$ denote the set of vector fields on P. We distinguish between two maps - $\phi^p(\cdot) : G \to P$ and $\phi_g(\cdot) : P \to P$ that are associated with $\phi(\cdot,\cdot)$ as follows:

$$\phi_g(p) \triangleq \phi(g,p) \quad \forall g \in G \quad \text{and} \quad \phi^p(g) \triangleq \phi(g,p) \quad \forall p \in P.$$

The ϕ - action is said to be *proper* if the map, $F : G \times P \to P \times P$ defined by

$$(g,p) \to (g \cdot p, p)$$

is proper. That is, the pre-image of F of every compact set is compact.

Assumption 7.1 *We assume that $\phi : G \times P \to P$ is a constant rank k proper left action.*

Definition 7.2. $\mathscr{O}(p)$, the orbit of ϕ through p is defined to be the set of points

$$\mathscr{O}(p) \triangleq \{\phi_g(p) \mid \forall\, g \in G\}.$$

Since the orbits are equivalence classes we will also denote it compactly as $[p]$. Let P/G denote the space of all orbits of ϕ with $\pi_\phi : P \to P/G$ denoting the canonical projection map. That is let $\pi_\phi(p) = [p]$. In what follows we summarize several smoothness properties of the orbits, $\mathscr{O}(p)$, and the collection of orbits M/G.

Define for every $\zeta \in \mathscr{G}$, *the infinitesimal generator*, the vector field $\zeta_P \in \mathscr{X}_P$ that is explicitly given by

$$\zeta_P(p) \triangleq \frac{d}{dt}\Big|_{t=0} \phi_{\exp \zeta t}(p) \quad \forall \quad p \in P\,.$$

Note that the flow of $\zeta_P \in \mathscr{X}_P$ is thus $\Phi^t_{\zeta_P}(p) \triangleq \phi_{\exp\zeta t}(p)$, where $\exp \zeta t \in G \quad \forall t \in \mathbb{R}$. Consider the map $\phi^p : G \to P$ at the identity of the group. Then the associated tangent map $T_e\phi^p \cdot \zeta = \zeta_P(p)$. Since $T_e\phi^p$ is linear we have that

$$(\zeta + \eta)_P = \zeta_P + \eta_P,$$
$$(\alpha\,\zeta)_P = \alpha\,\zeta_P.$$

Note that by definition

$$(\mathrm{Ad}_g \zeta)_P(p) = T_e\phi^p \cdot \mathrm{Ad}_g \zeta.$$

Thus for left invariant actions

$$\begin{aligned}
(\mathrm{Ad}_g \zeta)_P(\phi_g(p)) &= T_e\phi^{\phi_g(p)} \cdot \mathrm{Ad}_g \zeta \\
&= \frac{d}{ds}\Big|_{s=0} \phi(g \exp(\zeta s) g^{-1}, \phi(g,p)) = \frac{d}{ds}\Big|_{s=0} \phi(g \exp(\zeta s), p) \\
&= \frac{d}{ds}\Big|_{s=0} \phi_g \circ \phi_{\exp(\zeta s)}(p) = T_p\phi_g \cdot T_e\phi^p \cdot \zeta
\end{aligned}$$

and hence that

$$T_p\phi_g \cdot \zeta_P(p) = T_p\phi_g \cdot (T_e\phi^p \cdot \zeta) = T_e\phi^{\phi_g(p)} \cdot \mathrm{Ad}_g\zeta = (\mathrm{Ad}_g\zeta)_P(\phi_g(p)) \neq \zeta_P(\phi_g(p)).$$

Differentiating this expression (see Appendix) it also follows that $[\eta_P, \zeta_P] = -[\eta, \zeta]_P$. Thus for left actions the assignment $\zeta \to \zeta_P$ is a Lie algebra antimorphism and the subspace of vectorfields $\mathscr{X}_G \triangleq \{\zeta_P : \zeta \in \mathscr{G}\}$ is a Lie-subalgebra of the space of vectorfields $\mathscr{X}$ on P. Since this distribution is involutive, it is integrable. Since it is tangent to the orbits at every point of the orbit these integral manifolds are in fact the orbits of the action. A similar equivalent result holds for right invariant actions as well.

———-

Let $\phi : G \times M \to M$ be a left action of G on M. Define for $\zeta \in \mathscr{G}$ the vector field $\zeta_M \in \mathscr{X}$ that is give by

$$\zeta_M(m) \triangleq \frac{d}{dt}\Big|_{t=0}\phi_{\exp\zeta t}(m) \quad \forall \quad m \in M\,.$$

We note that

$$\begin{aligned}(\mathrm{Ad}_g\zeta)_M(m) &= \frac{d}{dt}\Big|_{t=0}\,\phi_{\exp t\mathrm{Ad}_g\zeta}(m) = \frac{d}{dt}\Big|_{t=0}\,\phi_{I_g(\exp t\zeta)}(m) = \frac{d}{dt}\Big|_{t=0}\,\phi_g \circ \phi_{\exp t\zeta} \circ \phi_{g^{-1}}(m)\\ &= \frac{d}{dt}\Big|_{t=0}\,\phi_g \circ \phi_{\exp t\zeta}(\phi_{g^{-1}}(m)) = (T_{\phi_{g^{-1}}(m)}\phi_g)\zeta_M(\phi_{g^{-1}}(m)) = (\phi^*_{g^{-1}}\zeta_M)(m).\end{aligned}$$

Since $\phi_{\exp\zeta t}(m) \triangleq \Phi_t^{\zeta_M}$ is the flow of ζ_M by setting $g = \exp(t\eta)$ in the above expression we have

$$\begin{aligned}(\mathrm{Ad}_{\exp(t\eta)}\zeta)_M(m) &= (T_{\phi_{\exp(-t\eta)}(m)}\phi_{\exp(t\eta)})\zeta_M(\phi_{\exp(-t\eta)}(m)) = \phi^*_{\exp(-t\eta)}\zeta_M(m)\\ &= (\Phi_t^{-\eta_M})^*\zeta_M(m).\end{aligned}$$

Differentiating with respect to t at $t = 0$ we have

$$(\mathrm{ad}_\eta\zeta)_M = -[\eta_M, \zeta_M]. \tag{7.110}$$

Note that if $c(t) \in G$ is a smooth curve such that $c(0) = e$ and $\dot{c}(0) = \zeta$ then by definition $\frac{d}{dt}|_{t=0}\phi_{c(t)}(m) = \zeta_M(m)$. For some fixed $m \in M$ consider the map $\phi^m : G \to M$ that is given by $\phi^m(g) \triangleq \phi_g(m)$. Then $T_e\phi^m \cdot \zeta = \zeta_M$. Thus since $T_e\phi^m$ is linear we have that $(\zeta + \eta)_M = \zeta_M + \eta_M$ and $(\alpha\,\zeta)_M = \alpha\,\zeta_M$.

7.9 Poisson Manifolds and Hamiltonian Dynamcs

Let $\mathscr{F}(M)$ denote the space of infinitely differentiable functions on a smooth n-dimensional manifold M. A bilinear operators $\{\cdot,\cdot\} : \mathscr{F} \times \mathscr{F} \to \mathbb{R}$ such that $\mathscr{F}$ is a Lie algebra with Lie bracket $\{\cdot,\cdot\}$ and $\{\cdot,\cdot\}$ with each argument fixed is a derivation on $\mathscr{F}$. That is

$$\{fg,h\} = f\{g,h\} + g\{f,h\} \quad \forall \quad f,g \in \mathscr{F}$$

Let $\mathscr{X}(M)$ denote the space of infinitely differentiable vector fields on M. A vector field $X_H \in \mathscr{X}$ is said to be Hamiltonian with respect to $H \in \mathscr{F}$ if

$$X_H(f) = \{f,H\} \quad \forall \quad f \in \mathscr{F}.$$

Let G be a Lie group and $\phi : G \times M \to M$ be a left action of G on M. The action is said to be canonical if

$$\phi_g^*\{f,g\} = \{\phi_g^* f, \phi_g^* g\} \quad \forall \quad f,g \in \mathscr{F}.$$

Define for $\zeta \in \mathscr{G}$ the vector field $\zeta_M \in \mathscr{X}$ that is give by

$$\zeta_M(m) \triangleq \frac{d}{dt}\Big|_{t=0} \phi_{\exp \zeta t}(m) \quad \forall \quad m \in M\,.$$

Define the map $J : \mathscr{G} \to \mathscr{F}$ such that for each $\zeta \in \mathscr{G}$ the vector field ζ_M is Hamiltonian with respect to J_ζ. That is

$$X_{J_\zeta} = \zeta_M.$$

or alternatively

$$\zeta_M(f) = \{f, J_\zeta\} \quad \forall \quad f \in \mathscr{F}.$$

Note that J_ζ is unique only upto an additive Casimir function and is thus a linear function. The map $\mathbf{J} : M \to \mathscr{G}^*$ defined such that

$$\langle \mathbf{J}(m), \zeta \rangle \triangleq J_\zeta(m),$$

is called a *momentum map* of the action ϕ.

If $\zeta_M(H) = 0$ for all $\zeta \in \mathscr{G}$ then $\{H, J_\zeta\} = 0$ for all $\zeta \in \mathscr{G}$. Hence $X_H(J_\zeta)$ for all $\zeta \in \mathscr{G}$. Thus $J_\zeta \in \mathscr{F}$ is conserved along the flow of X_H for all $J \in \mathscr{G}$ and hence $\mathbf{J}$ is conserved along the flow of X_H. This is Noether's theorem in the Hamiltonian setting.

Theorem 7.8. Noether's Theorem: *If $\phi : G \times M \to M$ is a canonical left action of G on M and $H \in \mathscr{F}$ is invariant under the action of ϕ then the momentum map $\mathbf{J}$ corresponding to the action of ϕ is conserved under the flow of X_H.*

7.10 Affine Connection on Bundles

A *Fibre bundle* E over a *base manifold* M with *fibre* manifold F is a 5-tuple $\{E, M, \pi, F\}$ such that

$$\pi : E \to M$$

is onto, and if $\{\mathscr{U}, \mathscr{V}, \cdots\}$ is a covering of M then there exist associated isomorphisms

$$\phi_U : \mathscr{U} \times F \mapsto \pi^{-1}(\mathscr{U}),$$

such that if $p \in M$, $s \in \pi^{-1}(p)$, and $\phi_U^{-1}(s) = (p, \psi^U)$ then in overlapping charts $\psi^V = c_{VU}\psi^U$ where $c_{VU} \in \mathscr{D}(F)$, $c_{UV} = c_{VU}^{-1}$, and $c_{UW} = c_{UV}c_{VW}$. Here $\mathscr{D}(F)$ is the set of all diffeomorphisms of F and the c_{VU} are called the *transition maps*. What this says is that each fibre $\pi^{-1}(p)$ is isomorphic to a copy of the fibre F. A section of the bundle E is a differentiable map $\Psi : M \to E$ such that $\pi \circ \Psi = id$. When all $c_{VU} \in G \subset \mathscr{D}(F)$ for some Lie group we call G the *structure group* of the bundle.

When F is a k-dimensional vector space it is diffeomorphic to $\mathbb{R}^k$ or $\mathbb{C}^k$ and we call the bundle a *rank-k vector bundle*. If the fibre F is the same as the structure group G itself and the transition maps, c_{VU}, act on the left of $F = G$ we call the bundle a *principle bundle*.

7.10.1 Connection in a Vector Bundle

In the following we will first restrict our attention to vector bundles. Let $e = \{e_1, \cdots, e_k\}$ be a frame of V^k. Define the frame of sections $\mathbf{e}^U = \{\mathbf{e}_1^U, \cdots, \mathbf{e}_k^U\}$ over $\mathscr{U}$ where $\mathbf{e}_i^U(p) \triangleq \phi^U(p, e_i)$. Then a general section $\psi : M \mapsto E$ can be locally expressed as $(p, \mathbf{e}^U\psi_U)$ where $\psi_V = c_{VU}\psi_U$.

A connection in the vector bundle E is a linear map ∇ such that it assigns to a section ψ of E an E valued 1-form such that the following Leibnitz rule holds

$$\nabla \psi f = (\nabla \psi) f + \psi \otimes df \tag{7.111}$$

for a function f on M. Note that then $\nabla_X \psi f = (\nabla_X \psi) f + \psi df(X) = (\nabla_X \psi) f + \psi \mathscr{L}_X f$.

For convenience of notation we will drop the superscript that represents the coordinate patch considered and denote by $\mathbf{e} = \{\mathbf{e}_1, \cdots, \mathbf{e}_k\}$ a frame of sections. Then in this frame we can write $\mathbf{e}\psi = \mathbf{e}_i\psi^i$. Then we can write

$$\nabla \mathbf{e}\psi = \nabla \mathbf{e}_i \psi^i = (\nabla \mathbf{e}_i)\psi^i + \mathbf{e}_i \otimes d\psi^i.$$

Let us define

$$\nabla \mathbf{e}_i \triangleq \mathbf{e}_j \otimes \omega_i^j,$$

where ω_i^j are 1-forms on M. Then we may write the above in components as

$$\nabla \mathbf{e}\psi = \mathbf{e}_j \otimes (\omega_i^j \psi^i + d\psi^j) = \mathbf{e} \otimes (\omega\psi + d\psi), \tag{7.112}$$

where ω is a matrix of 1-forms ω_i^j. The right hand side motivates us to define

$$\nabla \psi \triangleq d\psi + \omega\psi. \tag{7.113}$$

From For a vector field X on M and a section ψ of the bundle we have

$$\nabla_X \psi = d\psi(X) + \omega(X)\psi = \mathscr{L}_X \psi + \omega(X)\psi, \tag{7.114}$$

where the Lie derivative $\mathscr{L}$ here is to be taken component wise. That is $(\mathscr{L}_X \psi)^i = \mathscr{L}_X \psi^i$. Thus we also have

$$\nabla_{\partial_i} = \partial_i + \omega(\partial_i). \tag{7.115}$$

In two different patches

$$\mathbf{e}^U \otimes (\omega_U \psi_U + d\psi_U) = \mathbf{e}^V \otimes (\omega_V \psi_V + d\psi_V). \tag{7.116}$$

We know that $\psi_V = c_{VU}\psi_U$ and $\mathbf{e}^V = \mathbf{e}^U c_{UV}$. Thus

$$\mathbf{e}^U \otimes (\omega_U c_{UV} \psi_V + d(c_{UV}\psi_V) = \mathbf{e}^U c_{UV} \otimes (\omega_V \psi_V + d\psi_V). \tag{7.117}$$

From which we have

$$\begin{aligned} c_{UV}(\omega_V \psi_V + d\psi_V) &= (\omega_U c_{UV} \psi_V + d(c_{UV}\psi_V), \\ &= (\omega_U c_{UV} \psi_V + (dc_{UV})\psi_V + c_{UV} d\psi_V \end{aligned}$$

and hence

$$(\omega_V \psi_V + d\psi_V) = (c_{UV}^{-1}\omega_U c_{UV} + c_{UV}^{-1}(dc_{UV}))\psi_V + d\psi_V.$$

From which we have that if the locally defined matrix of 1-forms $\{\omega_U\}$ define a connection then in an overlap they are related by

$$\omega_V = c_{UV}^{-1}\omega_U c_{UV} + c_{UV}^{-1} dc_{UV}, \tag{7.118}$$

and that if $\{\psi_U\}$ is a section of the bundle then

$$\nabla \psi_V = c_{VU} \nabla \psi_U, \tag{7.119}$$

as desired.

7.10.2 Parallel Transport

Let X be a smooth vector filed on M and denote by ϕ_X^t the flow generated by X. We will say that a section ϑ is parallel along the flow generated by X if $\nabla_X \vartheta = 0$. That is

$$\nabla_X \vartheta = d\vartheta(X) + \omega(X)\vartheta = \frac{d\vartheta}{dt} + \omega(X)\vartheta = 0.$$

Hence if ϑ is parallel along the flow of X then ϑ must satisfy the differential equation

$$\frac{d\vartheta}{dt} = -\omega(X)\vartheta. \tag{7.120}$$

From which we have that if ϑ is parallel along X then

$$\vartheta(\phi^t(p)) \approx \vartheta(p) - t\omega(X(p))\vartheta(p), \tag{7.121}$$

for small t. Denote by $\mathscr{T}_{\phi^t(p)}(\vartheta(p))$ the parallel transport of $\vartheta(p)$ at the fibre over p to the fibre over $\phi^t(p)$. Then from above we see that

$$\mathscr{T}_{\phi^t(p)}(\vartheta(p)) \approx \vartheta(p) - t\omega(X(p))\vartheta(p), \tag{7.122}$$

for small t.

Let ψ be a section of the bundle then

$$\begin{aligned}\frac{\psi(\phi^t(p)) - \mathscr{T}_{\phi^t(p)}(\psi(p))}{t} &\approx \frac{\psi(\phi^t(p)) - \psi(p) + t\omega(X(p))\psi(p)}{t},\\ &\approx \frac{\psi(\phi^t(p)) - \psi(p)}{t} + \omega(X(p))\psi(p).\end{aligned}$$

Hence we see that

$$\begin{aligned}\lim_{t\to 0}\left(\frac{\psi(\phi^t(p)) - \mathscr{T}_{\phi^t(p)}(\psi(p))}{t}\right) &= \lim_{t\to 0}\left(\frac{\psi(\phi^t(p)) - \psi(p)}{t}\right) + \omega(X(p))\psi(p)\\ &= \mathscr{L}_X\psi + \omega(X(p))\psi(p)\\ &= \nabla_X\psi, \end{aligned}\tag{7.123}$$

and hence that $\nabla_X\psi$ measures the infinitesimal difference between ψ and its parallel transport along the integral curves of X.

7.10.3 Curvature

We may also compute

$$\begin{aligned}\nabla_X\nabla_Y\psi &= \nabla_X(\mathscr{L}_Y\psi + (i_Y\omega)\psi) = \mathscr{L}_X(\mathscr{L}_Y\psi + (i_Y\omega)\psi) + (i_X\omega)(\mathscr{L}_Y\psi + (i_Y\omega)\psi)\\ &= \mathscr{L}_X\mathscr{L}_Y\psi + \mathscr{L}_X(i_Y\omega\psi) + (i_X\omega)\mathscr{L}_Y\psi + (i_X\omega)(i_Y\omega)\psi\\ &= \mathscr{L}_X\mathscr{L}_Y\psi + (i_Xdi_Y\omega)\psi + (i_Y\omega)\mathscr{L}_X\psi + (i_X\omega)\mathscr{L}_Y\psi + (i_X\omega)(i_Y\omega)\psi\end{aligned}\tag{7.124}$$

and thus similarly we also have

$$\nabla_Y\nabla_X\psi = \mathscr{L}_Y\mathscr{L}_X\psi + (i_Ydi_X\omega)\psi + (i_X\omega)\mathscr{L}_Y\psi + (i_Y\omega)\mathscr{L}_X\psi + (i_Y\omega)(i_X\omega)\psi.$$

These second order derivatives are not tensors.Taking the difference between them we have

$$\nabla_X\nabla_Y\psi - \nabla_Y\nabla_X\psi = (\mathscr{L}_X\mathscr{L}_Y - \mathscr{L}_Y\mathscr{L}_X)\psi + ((i_Xdi_Y - i_Ydi_X)\omega)\psi + i_Yi_X(\omega\wedge\omega)\,\psi.$$

We see that

$$\begin{aligned}
i_X d i_Y - i_Y d i_X = (\mathscr{L}_X - d i_X) i_Y - i_Y d i_X &= \mathscr{L}_X i_Y - d i_X i_Y - i_Y d i_X \\
&= i_Y \mathscr{L}_X + i_{[X,Y]} - d i_X i_Y - i_Y d i_X \\
&= i_Y d i_X + i_Y i_X d + i_{[X,Y]} - d i_X i_Y - i_Y d i_X \\
&= i_Y i_X d + i_{[X,Y]} + d i_Y i_X,
\end{aligned}$$

and $(\mathscr{L}_X \mathscr{L}_Y - \mathscr{L}_Y \mathscr{L}_X)\psi^i = \mathscr{L}_{[X,Y]}\psi^i$. Hence that

$$\begin{aligned}
\nabla_X \nabla_Y \psi - \nabla_Y \nabla_X \psi &= \mathscr{L}_{[X,Y]}\psi + i_{[X,Y]}\omega\psi + i_Y i_X d\omega\psi + i_Y i_X (\omega \wedge \omega)\psi \\
&= (\mathscr{L}_{[X,Y]}\psi + i_{[X,Y]}\omega\psi) + i_Y i_X (d\omega + (\omega \wedge \omega))\psi \\
&= \nabla_{[X,Y]}\psi + \theta(X,Y)\psi.
\end{aligned}$$

The 2-form

$$\theta = d\omega + (\omega \wedge \omega), \tag{7.125}$$

is called the curvature 2-form and the tensor

$$R(X,Y)\psi \triangleq \theta(X,Y)\psi = \nabla_X \nabla_Y \psi - \nabla_Y \nabla_X \psi - \nabla_{[X,Y]}\psi. \tag{7.126}$$

is called the curvature tensor. Let $R(e_i,e_j)e_s = e_r R^r_{sij}$. From this and (7.157) we see that $R^r_{sji} = -R^r_{sij}$ (note that r and s are fibre indices while i and j are manifold indices). Then we can write

$$R(e_i,e_j)e_s = e_r \left(\sum_{ij} \frac{1}{2} R^r_{sij} \sigma_i \wedge \sigma_j (e_i,e_j) \right).$$

Thus since we may also write $\theta^r_s = \sum_{i<j} \theta^r_{sij} \sigma_i \wedge \sigma_j$ we see that $\theta^r_{sij} = \frac{1}{2} R^r_{sij}$.

Observe that from (7.125) we have that

$$\begin{aligned}
d\theta &= d\omega \wedge \omega - \omega \wedge d\omega = (\theta - \omega \wedge \omega) \wedge \omega - \omega \wedge (\theta - \omega \wedge \omega) \\
&= \theta \wedge \omega - \omega \wedge \theta.
\end{aligned} \tag{7.127}$$

This is called the *second Bianchi* identity.

If $\{\theta^U\}$ is a collection of vector valued connection 2-forms we see from (7.118) and noting that since $c_{UV}^{-1} c_{UV} = id$ that $dc_{UV}^{-1} c_{UV} + c_{UV}^{-1} dc_{UV} = 0$

$$\theta_V = c_{UV}^{-1} \theta_U c_{UV}. \tag{7.128}$$

7.10.4 $SO(3)$ *and* $SU(2)$

*The contents of this section are beyond the scope of these lecture notes. I have included them here for my understanding and completeness and also to motivate you to pursue this beautiful subject further. What you need to take out of this section are the two formulas (**??**) and (**??**) that allow you to do rigid body simulations without having to resort to singular co-ordinates such as Euler angles.*

Let $\mathrm{SU}(2)=\{A\in \mathrm{GL}(2,\mathbb{C})\;|\;\langle Ax,Ay\rangle=\langle x,y\rangle \;\text{ and }\; \det(A)=1\}$ and $\mathrm{su}(2)$ be the Lie algebra of $\mathrm{SU}(2)$ that consists of the traceless skew-hermitian matrices. The Pauli spin matrices

$$\sigma_1=\begin{bmatrix}0&1\\1&0\end{bmatrix},\;\;\sigma_2=\begin{bmatrix}0&-i\\i&0\end{bmatrix},\;\;\sigma_3=\begin{bmatrix}1&0\\0&-1\end{bmatrix},$$

form a basis for the space of 2×2 hermitian matrices. Then it can be shown that $\{\tilde{e}_1,\tilde{e}_2,\tilde{e}_3\}=\{\frac{1}{2i}\sigma_1,\frac{1}{2i}\sigma_2,\frac{1}{2i}\sigma_3\}$ form a basis for $\mathrm{su}(2)$.

Define the isomorphism $\tilde{}:\mathbb{R}^3\to\mathrm{su}(2)$ by the relationship

$$\tilde{x}=\frac{1}{2i}(x_1\sigma_1+x_2\sigma_2+x_3\sigma_3)\triangleq\frac{1}{2i}(x\cdot\sigma)=\frac{1}{2}\begin{bmatrix}-ix_3 & -ix_1-x_2\\ -ix_1+x_2 & ix_3\end{bmatrix},$$

for $x=(x_1,x_2,x_3)\in\mathbb{R}^3$. Let $\{e_1,e_2,e_3\}$ be the standard basis on $\mathbb{R}^3$. Then one can show that

$$\tilde{e}_i\tilde{e}_j=\frac{1}{2}\tilde{e}_k \tag{7.129}$$

$$\tilde{e}_j\tilde{e}_i=-\frac{1}{2}\tilde{e}_k \tag{7.130}$$

where (i,j,k) is a cyclic permutation of $(1,2,3)$ and

$$\tilde{e}_i\tilde{e}_i=\frac{1}{4}I_{2\times 2}. \tag{7.131}$$

Thus it can be shown that

$$x\cdot y=-2\mathrm{tr}(\tilde{x}\tilde{y}). \tag{7.132}$$

It also follows from (7.130) that

$$[\tilde{e}_i,\tilde{e}_j]=\tilde{e}_k, \tag{7.133}$$

and hence that

$$[\tilde{x},\tilde{y}]=\widetilde{(x\times y)}. \tag{7.134}$$

(or alternatively $\mathrm{ad}_{\tilde{x}}\tilde{y}=\widetilde{\hat{x}y}=\widetilde{\mathrm{ad}_x y}$). This implies that the map $\tilde{}:\mathbb{R}^3\to\mathrm{su}(2)$ is a Lie algebra isomorphism.

Consider the adjoint action $\mathrm{Ad}:\mathrm{SU}(2)\times\mathrm{su}(2)\to\mathrm{su}(2)$ given by

$$\mathrm{Ad}_u\tilde{x} = u\tilde{x}u^\dagger$$

Thus we can also define the map $\mathrm{Ad} : \mathrm{SU}(2) \times \mathbb{R}^3 \to \mathbb{R}^3$ by the identification

$$\widetilde{\mathrm{Ad}_u x} \triangleq \mathrm{Ad}_u\tilde{x} = u\left(\frac{1}{2i}(x\cdot\sigma)\right)u^\dagger,$$

where $\mathrm{Ad}_u \in \mathrm{GL}(3,\mathbb{R})$. This is also a group action since it is easily verifiable that $\mathrm{Ad}_{uv}x = \mathrm{Ad}_u\mathrm{Ad}_v x$ and $\mathrm{Ad}_I x = x$ for all $x \in \mathbb{R}^3$.

Since

$$\mathrm{Ad}_u x\cdot \mathrm{Ad}_u x = -2\mathrm{trace}(\tilde{x}\tilde{x}) = ||x||^2$$

we see that the group homomorphism is in fact $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{O}(3)$. Since $\mathrm{SU}(2)$ is connected any u can be path connected to I. Let $u(t)$ be such a path such that $u(0) = u$ and $u(1) = I$. Then since $\det(\mathrm{Ad}_{u(t)}) = \pm 1$ for all t and since det is a continuous map and $\det(\mathrm{Ad}_{u(1)}) = 1$ we have that $\det(\mathrm{Ad}_{u(1)}) = 1$ for all t and in particular $\det(\mathrm{Ad}_u) = \det(\mathrm{Ad}_{u(0)}) = 1$. This implies that $\mathrm{Ad}(\mathrm{SU}(2)) \subset \mathrm{SO}(3)$.

We also see that $\mathrm{Ad}_{-I}x = x$ for all $x \in \mathbb{R}^3$. Since $\mathrm{Ad}_u x = \mathrm{Ad}_v x$, implies $\mathrm{Ad}_{v^\dagger u}x = x$, we have that $v = \pm u$. Which implies that this group homomorphism is $2:1$. The isotropy subgroup of the action for any x is seen to be given by the subgroup $H = \{I, -I\} = \mathrm{Ad}^{-1}(I)$. Thus if this representation is onto then it is an elementary result of group theory that $\mathrm{SU}(2)/\{I,-I\} \simeq \mathrm{SO}(3)$.

To show that the Ad representation is onto let us consider the 1-parameter subgroup $u(t) = \exp(\tilde{\Omega}t)$ and

$$\widetilde{\mathrm{Ad}_{\exp(\tilde{\Omega}t)}x} = \mathrm{Ad}_{\exp(\tilde{\Omega}t)}\tilde{x} = \exp\left(t\mathrm{ad}_{\tilde{\Omega}}\right)\tilde{x} = \widetilde{\exp t\left(\mathrm{ad}_{\Omega}\right)x} = \widetilde{\mathrm{Ad}_{\exp(\widehat{\Omega}t)}x} = \widetilde{\exp t\left(\widehat{\Omega}\right)x}.$$

Here the third equality follows from the fact

$$\exp\left(t\mathrm{ad}_{\tilde{\Omega}}\right)\tilde{x} = \tilde{x} + t\mathrm{ad}_{\tilde{\Omega}}\tilde{x} + \frac{1}{2}t^2\mathrm{ad}^2_{\tilde{\Omega}}\tilde{x} + \mathrm{H.O.T.} = \tilde{x} + t\widetilde{\mathrm{ad}_{\Omega}x} + \frac{1}{2}t^2\widetilde{\mathrm{ad}^2_{\Omega}x} + \mathrm{H.O.T.}$$

Thus

$$\mathrm{Ad}_{\exp(\tilde{\Omega}t)}x = \exp\left(t\,\widehat{\Omega}\right)x$$

and

$$\mathrm{Ad}_{\exp(\tilde{\Omega}t)} = \exp\left(t\,\widehat{\Omega}\right).$$

Since given any $R \in \mathrm{SO}(3)$ can be expressed as $R = \exp\widehat{\Omega}$ we see that the Ad map is onto.

What this means is that given $u \in \mathrm{SU}(2)$ the explicit formula for $\mathrm{Ad}_u \in \mathrm{SO}(3)$ can be expressed as follows. First one solves $\tilde{\Omega}$ for $u = \exp\left(\tilde{\Omega}\right)$. This is always possible since the exp map is onto for compact Lie groups. Then we have that $\mathrm{Ad}_u = \exp\left(\widehat{\Omega}\right)$. We have seen that this $\mathrm{Ad} : SU(2) \to SO(3)$ is 2:1 where $\mathrm{Ad}_u = \mathrm{Ad}_{-u}$. Thus we have that $\pm u$ can be identified with $\exp\left(\widehat{\Omega}\right) \in SO(3)$.

Below we proceed to find explicit expressions for the exp map on $su(2)$ and $so(3)$ respectively in order to explicitly construct this identification.

7.10.5 The exp map on $\mathrm{su}(2)$

Let $\Omega \in \mathbb{R}^3$, $\theta = ||\Omega||$ and $n = \Omega/||\Omega||$. Then it can be shown using (7.129)–(7.131) that

$$\exp\left(\tilde{\Omega}\right) = \exp\left(\theta \frac{1}{2i}(n\cdot\sigma)\right) = \left(\cos\frac{\theta}{2}\right) I_{2\times 2} - i\left(\sin\frac{\theta}{2}\right)(n\cdot\sigma). \tag{7.135}$$

Notice that (7.135) can be written as

$$\exp\tilde{\Omega} = q_0 I_{2\times 2} - i(w\cdot\sigma) \tag{7.136}$$

where $q_0 = \cos\left(\frac{||\Omega||}{2}\right)$ and $w = \sin\left(\frac{||\Omega||}{2}\right)\frac{\Omega}{||\Omega||}$. Thus we see that $q_0^2 + ||w||^2 = 1$ and hence (7.136) says that $q = (q_0, w) \in \mathbb{S}^3 = \{x \in \mathbb{R}^4 \mid ||x|| = 1\}$ can be identified with a $u \in \mathrm{SU}(2)$ in the following way:

$$q = (q_0, w_1, w_2, w_3) \in \mathbb{S}^3 \to \begin{bmatrix} q_0 - iw_3 & -w_2 - iw_1 \\ w_2 - iw_1 & q_0 + iw_3 \end{bmatrix} \in \mathrm{SU}(2).$$

It is also easy to see that in this same fashion a given $u \in \mathrm{SU}(2)$ can be uniquely identified with a $q = (q_0, w) \in \mathbb{S}^3$. Specifically this identification shows that $SU(2) \simeq \mathbb{S}^3$.

Since $\mathrm{SU}(2)$ is compact the exp map is onto. Thus any $u \in \mathrm{SU}(2)$ can be written as $u = \exp\tilde{\Omega}$ for some $\Omega \in \mathbb{R}^3$ as well. From the above identifications we then see that this Ω is given by $\Omega = \theta\frac{w}{||w||}$ where $\cos\frac{\theta}{2} = q_0$ and $q = (q_0, w) \in \mathbb{S}^3$ is the unit quaternion that corresponds to the given u.

In summary we have shown that given a $u \in SU(2)$ there exists a corresponding unique $(q_0, w) \in \mathbb{S}^3$ given by $u = q_0 - i(w\cdot\sigma) = \exp\left(\tilde{\Omega}\right)$ where $\Omega = \theta\frac{w}{||w||}$ and $\theta = 2\arccos(q_0)$. In the previous section we saw that u can be identified with $\exp(\widehat{\Omega}) \in SO(3)$. Thus any given $u \simeq (q_0, w)$ can be made to correspond to a rotation about the axis w by an angle equal to $\theta = 2\arccos(q_0)$.

7.10.5.1 The exp map on $\mathrm{so}(3)$

Let $\Omega \in \mathbb{R}^3$ then $\exp\left(\widehat{\Omega}\right)$ is a rotation about Ω by an angle $\theta = ||\Omega||$. One can show using the properties of cross product proven in exercise-**??** that

$$\widehat{\Omega}^2 w = (\Omega\cdot w)\Omega - ||\Omega||^2 w = (\Omega\Omega^T - ||\Omega||^2 I)w$$

hence that

$$\widehat{\Omega}^3 = -||\Omega||^2\widehat{\Omega}, \quad \widehat{\Omega}^4 = -||\Omega||^2\widehat{\Omega}^2, \quad \widehat{\Omega}^5 = ||\Omega||^4\widehat{\Omega}^2, \quad \cdots$$

From this we have the Rodrigues formula

$$\exp\left(\widehat{\Omega}\right) = I + \frac{\sin||\Omega||}{||\Omega||}\widehat{\Omega} + \frac{1}{2}\left(\frac{\sin\frac{||\Omega||}{2}}{\frac{||\Omega||}{2}}\right)^2\widehat{\Omega}^2, \tag{7.137}$$

$$= I + \sin\theta\widehat{n} + (1-\cos\theta)\widehat{n}^2 \tag{7.138}$$

$$= I + \sin\theta\widehat{n} + (1-\cos\theta)[nn^T - I] \tag{7.139}$$

where we have set $\theta = ||\Omega||$ and $n = \Omega/||\Omega||$ in the last equality. Let $w = \sin\left(\frac{\theta}{2}\right)n$. Then re-arranging the expression (7.138) we have

$$\begin{aligned} R = \exp\left(\widehat{\Omega}\right) &= I + \sin\theta\widehat{n} + (1-\cos\theta)\widehat{n}^2 \\ &= I + 2\sin(\theta/2)\cos(\theta/2)\widehat{n} + 2\sin^2(\theta/2)\widehat{n}^2 \\ &= I + 2q_0\widehat{w} + 2\widehat{w}^2 \\ &= (q_0^2 - ||w||^2)I + 2q_0\widehat{w} + 2ww^T \end{aligned} \tag{7.140}$$

In summary, we have the following. The group of rotations SO(3) is diffeomorphic to $SU(2)/\{I,-I\}$ (which is in turn diffeomorphic to $\mathbb{S}^3/\{1,-1\}$). where the diffeomorphism is explicitly given by the Rodrigues formula

$$R = I + 2q_0\widehat{w} + 2\widehat{w}^2 \tag{7.141}$$

for $q = (q_0, w) \in \mathbb{S}^3$. Here R corresponds to a rotation about w by an angle of $\theta = 2\cos^{-1}(q_0)$. Using algebraic manipulations it can also be shown that the differential equation $\dot{R} = R\widehat{\Omega}$ then corresponds to

$$\begin{bmatrix} \dot{q}_0 \\ \dot{w} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -\Omega\cdot w \\ q_0\Omega - \Omega\times w \end{bmatrix}. \tag{7.142}$$

We conclude this section by noting that we can also find the quaternions $q = (q_0, w)$ that correspond to a given rotation matrix R using the following two expressions obtained from (7.140):

$$\operatorname{trace}(R) = 2\cos\theta + 1, \tag{7.143}$$

$$R - R^T = 4\cos\left(\frac{\theta}{2}\right)\widehat{w}. \tag{7.144}$$

The first expression determines $q_0 = \cos\frac{\theta}{2}$ and the second expression determines w.

7.10.6 Example: Tangent Bundles Over Riemannian Manifolds

On a Riemannian manifold on each open set $\mathscr{U}$ of M we can define a frame $\mathbf{e}_U = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$. If the connection is metric then in this frame

$$\begin{aligned}
\mathscr{L}_v\langle\langle \mathbf{e}_i,\mathbf{e}_j\rangle\rangle &= \langle\langle \nabla_v \mathbf{e}_i,\mathbf{e}_j\rangle\rangle + \langle\langle \mathbf{e}_i,\nabla_v \mathbf{e}_j\rangle\rangle \\
&= \langle\langle \mathbf{e}_s\omega^s_i(v),\mathbf{e}_j\rangle\rangle + \langle\langle \mathbf{e}_i,\mathbf{e}_s\omega^s_j(v)\rangle\rangle \\
&= g_{sj}\omega^s_i(v) + g_{is}\omega^s_j(v)
\end{aligned}$$

Define $\omega_{ij} \triangleq g_{is}\omega^s_j$ and $\theta_{ij} \triangleq g_{is}\theta^s_j$. Then the above reads

$$dg_{ij} = g_{sj}\omega^s_i + g_{is}\omega^s_j = \omega_{ji} + \omega_{ij}. \tag{7.145}$$

On a tangent bundle one can define the torsion of the connection to be

$$\tau(u,v) = \nabla_u v - \nabla_v u - [u,v]. \tag{7.146}$$

Note that this makes sense only on a tangent bundle. In terms of components we have

$$\begin{aligned}
\tau^k_{ij} &= \sigma^k(\tau(\mathbf{e}_i,\mathbf{e}_j)) = \sigma^k(\omega(\mathbf{e}_i)\mathbf{e}_j - \omega(\mathbf{e}_j)\mathbf{e}_i - [\mathbf{e}_i,\mathbf{e}_j]) \\
&= (\omega^k_{ij} - \omega^k_{ji}) - \sigma^k([\mathbf{e}_i,\mathbf{e}_j]).
\end{aligned} \tag{7.147}$$

Since using (7.19)–(7.21) it can be shown that $\sigma^k([\mathbf{e}_i,\mathbf{e}_j]) = -d\sigma^k(\mathbf{e}_i,\mathbf{e}_j)$ we have

$$d\sigma^k(\mathbf{e}_i,\mathbf{e}_j) = \tau^k_{ij} - (\omega^k_{ij} - \omega^k_{ji}).$$

Thus we see that

$$d\sigma^k = \sum_{ij}\left(-(\omega^k_{ij}\sigma^i)\wedge\sigma^j + \frac{1}{2}\tau^k_{ij}\sigma^i\wedge\sigma^j\right),$$

and hence that

$$d\sigma^k = -\omega^k_j\wedge\sigma^j + \tau^k, \tag{7.148}$$

where $\tau^k \triangleq \sum_{i<j}\tau^k_{ij}\sigma^i\wedge\sigma^j$. We can also write this in matrix form as

$$d\sigma = -\omega\wedge\sigma + \tau. \tag{7.149}$$

This relationship is called the *Cartan's structural equations*.

In an orthonormal frame $g_{ij} = \delta_{ij}$ hence we have that $\omega_{ij} \triangleq g_{ij}\omega^i_j = \omega^i_j$ and similarly $\theta_{ij} = \theta^i_j$ and

$$\omega_{ij} = -\omega_{ji}. \tag{7.150}$$

Similarly

$$\theta_{ij} = d\omega_{ij} + \omega_{is}\wedge\omega_{sj} = -d\omega_{ji} - \omega_{js}\wedge\omega_{si} = -\theta_{ji}. \tag{7.151}$$

Thus when e is an orthonormal frame the matrix of connection 1-forms ω and the matrix of curvature 2-forms θ are skew symmetric.

7.10.6.1 Tangent Bundle to Riemann Surfaces

Consider a Riemann surface S. Let $\mathbf{e} = \{\mathbf{e}_1, \mathbf{e}_2\}$ be an orthonormal basis in a patch $\mathscr{U}$. In this basis we see that $\omega_{11} = \omega_{22} = 0$ and $\omega_{12} = -\omega_{21}$. Thus we also see that $\theta_{11} = \theta_{22} = 0$ and

$$\theta_{12} = d\omega_{12} + \omega_{1s} \wedge \omega_{s2} = d\omega_{12} + \omega_{11} \wedge \omega_{12} + \omega_{12} \wedge \omega_{22} = d\omega_{12}.$$

7.10.6.2 Complex Line Bundles

Consider a complex line bundle E over a manifold M with fibre $\mathbb{C}$ and transition functions $c_{UV} : \mathbb{C} \to \mathbb{C}$.

Let $\mathbf{e} = (1,0)$ be the basis for $\mathbb{C}$. In a chart $\mathscr{U}$ on M any complex valued 1-form ω^c defines a connection on the bundle by the relationship

$$\nabla \mathbf{e}^U = \mathbf{e}^U \omega^c_U,$$

if in overlapping charts

$$\omega^c_V = c^{-1}_{UV} \omega^c_U c_{UV} + c^{-1}_{UV} dc_{UV} = \omega^c_U + c^{-1}_{UV} dc_{UV} = \omega^c_U + d \ln c_{UV}.$$

Let $\ln c_{UV} = f_{UV}$ then.

In Minkowski space R^4 consider the locally defined electromagnetic potential $\mathscr{A}$[2]. That is in a coordinate patch $\mathscr{U}$ define

$$\mathscr{A}_U = \phi_U dt + A_{U_1} dx^1 + A_{U_2} dx^2 + A_{U_3} dx^3.$$

and

$$\omega^c_U = \left(\frac{e}{i\hbar}\right) \mathscr{A}_U.$$

Let $\mathscr{V}$ be another chart that overlaps with $\mathscr{U}$ and let $\mathscr{A}_V$ be a 1-form defined on $\mathscr{V}$ where possibly $\mathscr{A}_U \neq \mathscr{A}_V$. However if we require that the electromagnetic 2-form $F = d\mathscr{A}$ then since F is well defined we must have that $d\mathscr{A}_U = d\mathscr{A}_V$. Thus if the overlap region is simply connected it follows that $\mathscr{A}_V - \mathscr{A}_U = df_{UV}$ for some real valued function in the overlap. Thus

$$\left(\frac{e}{i\hbar}\right) \mathscr{A}_V = \left(\frac{e}{i\hbar}\right) \mathscr{A}_U + d \ln c_{UV}.$$

and hence $\left(\frac{e}{i\hbar}\right) df_{UV} = d \ln c_{UV}$ and thus we must have that

$$c_{UV} = \exp\left(\frac{e}{i\hbar} f_{UV}\right).$$

One can easily show that $c_{VU} = c^{-1}_{UV}$ and $c_{UV} c_{VW} c_{WU} = 1$.

Thus we see that the electromagnetic field is a connection on a complex line bundle. We see that

[2] Note that this need not be globally defined.

$$\nabla_{\partial_t} = \frac{\partial}{\partial t} + \frac{e}{i\hbar}\phi, \tag{7.152}$$

$$\nabla_{\partial_\alpha} = \frac{\partial}{\partial x^\alpha} + \frac{e}{i\hbar}A_\alpha, \tag{7.153}$$

Let ψ be a section of the bundle E. Then the Schrodinger's equation can be written as

$$i\hbar\nabla_{\partial_t}\psi = \left(-\frac{\hbar^2}{2m}\right)\sum_{\alpha=1}^{3}\nabla_{\partial_\alpha}\nabla_{\partial_\alpha}\psi + V\psi \tag{7.154}$$

7.10.7 Covariant derivative of r-form sections

Note that for a section ψ of the vector bundle the covariant derivative operator can be considered as an operator that assigs a section ψ a fibre valued 1-form defined by $\nabla\psi = d\psi + \omega\psi$. This notion can be extended to r-form sections $\psi\otimes\alpha$ by requiring

$$\nabla(\psi\otimes\alpha) = (\nabla\psi)\wedge\alpha + \psi\otimes d\alpha. \tag{7.155}$$

Let M be a n-dimensional smooth manifold and let be $\{\mathscr{U},\mathscr{V},\cdots\}$ an open cover of M. Let $\{E,M,\pi,\mathbb{C}^N\}$ be a vector bundle with structure group G. Let $\mathscr{G}$ be the Lie algebra of G. Let $\{\omega_U\}$ be the locally defined connection 1-form on E. Let $\{e^U\}$ be a local section of frames on the bundle. Thus $e^V = e^U g_{UV}$ where $g_{UV}\in G$ are the transition maps. Let $\{\psi_U\}$ be a r-form section of the bundle E (that is r-forms that take values in the fibres $\mathbb{C}^N$). Then we have

$$\nabla e^U\psi_U = e^U d\psi_U + (\nabla e^U)\wedge\psi_U = e^U(d\psi_U + \omega^U\wedge\psi_U).$$

and we may define

$$\nabla\psi_U \triangleq d\psi_U + \omega^U\wedge\psi_U. \tag{7.156}$$

Then since

$$\begin{aligned}
\nabla\psi_U &= d(g_{UV}\psi_V) + (g_{VU}^{-1}\omega_V g_{VU} + g_{VU}^{-1}dg_{VU})\wedge(g_{UV}\psi_V)\\
&= g_{UV}(d\psi_V + \omega_V)\wedge\psi_V + (dg_{UV} + g_{VU}^{-1}dg_{VU}g_{UV})\wedge\psi_V\\
&= g_{UV}(d\psi_V + \omega_V)\wedge\psi_V\\
&= g_{UV}\nabla\psi_V
\end{aligned}$$

we see that (7.156) defines a covariant derivative on the r-form sections of the bundle E.

Let $q(t)\in M$ and if ψ is a r-form section such that $\nabla_{\dot q}\psi = 0$ along $q(t)$ then ψ is said to be *parallel* along $q(t)$ or it is said to be *covariant constant* along $q(t)$.

7.10.7.1 Curvature

Thus we define the curvature as the covariant derivative of the connection as follows:

$$\begin{aligned}\nabla\nabla\mathbf{e} &= \nabla(\mathbf{e}\omega) = (\nabla\mathbf{e})\wedge\omega + \mathbf{e}\otimes d\omega \\ &= \mathbf{e}\otimes(\omega\wedge\omega + d\omega) \\ &= \mathbf{e}\otimes\theta.\end{aligned} \tag{7.157}$$

Expanding (7.157) in components we have

$$\begin{aligned}\nabla\nabla\mathbf{e}_s &= \mathbf{e}_r\otimes(\omega^r_p\wedge\omega^p_s + d\omega^r_s), \\ \theta^r_s &= (\omega^r_p\wedge\omega^p_s + d\omega^r_s) \triangleq \frac{1}{2}R^r_{sij}dx^i\wedge dx^j.\end{aligned} \tag{7.158}$$

Consider a fibre valued r-form section ψ of the bundle. Then since $\nabla\mathbf{e}\otimes\psi = \mathbf{e}\otimes(\omega\wedge\psi + d\psi)$, we find

$$\begin{aligned}\nabla(\nabla\mathbf{e}\otimes\psi) &= \mathbf{e}\otimes(\omega\wedge\omega\wedge\psi + \omega\wedge d\psi + (d\omega\wedge\psi - \omega\wedge d\psi)) \\ &= \mathbf{e}\otimes(\omega\wedge\omega + d\omega)\wedge\psi \\ &= \mathbf{e}\otimes(\theta\wedge\psi).\end{aligned}$$

Thus we may define

$$\nabla\nabla\psi = \theta\wedge\psi \tag{7.159}$$

On a tangent bundle we may consider the vector valued 1-form $\mathbf{e}\otimes\sigma$. On a tangent bundle we have the Cartan's structure equations

$$d\sigma = -\omega\wedge\sigma + \tau,$$

where τ is the torsion 2-form. Then

$$\nabla\sigma = \omega\wedge\sigma + d\sigma = \tau,$$

and

$$\nabla\tau = \nabla\nabla\sigma = \theta\wedge\sigma.$$

For symmetric connections $\tau = 0$ and hence $\theta\wedge\sigma = 0$ and gives the *first Bianchi identity* valid only for symmetric connections:

$$\theta\wedge\sigma = \frac{1}{2}R^s_{rij}\sigma^i\wedge\sigma^j\wedge\sigma^r = 0,$$

which gives

$$R^s_{rij} + R^s_{ijr} + R^s_{jri} = 0.$$

7.11 Lie Algebra Valued p-forms

Let $\{E, M, \pi, \mathbb{C}^N\}$ be some vector bundle with structure group G. Let $\mathscr{G}$ be the Lie algebra of G. In this section we will consider $\mathscr{G}$ valued p-forms on M. Let $\mathbf{e} = \{\mathbf{e}_R\}$ be a basis for $\mathscr{G}$.

Let ϕ be a $\mathscr{G}$ valued globally defined p-form on M.

Then we may express

$$\phi = \mathbf{e}_R \otimes \phi^R,$$

where each ϕ^R is a p-form on M. The d-operator that takes $\mathscr{G}$ valued p-forms to $\mathscr{G}$ valued $p+1$-forms will then be given by

$$d\phi = \mathbf{e}_R \otimes d\phi^R.$$

The product between a $\mathscr{G}$ valued p-form ϕ and a $\mathscr{G}$ valued q-form ψ will be defined by

$$[\phi, \psi] \triangleq [E_R, E_S] \otimes \phi^R \wedge \psi^S$$

Note that

$$[\phi, \psi] = -[E_S, E_R] \otimes \phi^R \wedge \psi^S = (-1)^{pq+1}[E_S, E_R] \otimes \psi^S \wedge \phi^R = (-1)^{pq+1}[\psi, \phi].$$

Also we have

$$\begin{aligned} d[\phi, \psi] &= [E_R, E_S] \otimes d(\phi^R \wedge \psi^S) = [E_R, E_S] \otimes (d\phi^R \wedge \psi^S + (-1)^p \phi^R \wedge d\psi^S) \\ &= [d\phi, \psi] + (-1)^p [\phi, d\psi]. \end{aligned} \tag{7.160}$$

For matrix Lie groups we have

$$\begin{aligned} [\phi, \psi] &= [E_R, E_S] \otimes \phi^R \wedge \psi^S = E_R E_S \otimes \phi^R \wedge \psi^S - E_S E_R \otimes \phi^R \wedge \psi^S \\ &= E_R E_S \otimes \phi^R \wedge \psi^S - (-1)^{pq} E_S E_R \otimes \psi^S \wedge \phi^R \\ &= (\phi^R E_R) \wedge (\psi^S E_S) - (-1)^{pq} (\psi^S E_S) \wedge (\phi^R E_R) \\ &= \phi \wedge \psi - (-1)^{pq} \psi \wedge \phi \end{aligned} \tag{7.161}$$

Thus for matrix Lie groups if ω is a $\mathscr{G}$ valued 1-form we have that

$$[\omega, \omega] = \omega \wedge \omega - (-1)^1 \omega \wedge \omega = 2\omega \wedge \omega, \tag{7.162}$$

and that the curvature θ and $d\theta$ can then be written as

$$\theta = d\omega + \omega \wedge \omega = d\omega + \frac{1}{2}[\omega, \omega] \tag{7.163}$$

$$d\theta = \frac{1}{2}([d\omega, \omega] - [\omega, d\omega]) = [d\omega, \omega] = [\theta, \omega]. \tag{7.164}$$

7.11.1 Maure-Cartan Equations

The $\mathscr{G}$ valued globally defined 1-form on G referred to as the Maure-Cartan 1-form takes vectors at $g \in G$ to vectors at $e \in G$ by left translating the vector to the tangent plane at the identity. This is written down as

$$\Omega \triangleq E_R \otimes \sigma^R \triangleq g^{-1}dg. \tag{7.165}$$

We also see that for matrix Lie groups

$$d\Omega = dg^{-1} \wedge dg = -g^{-1}dgg^{-1} \wedge dg = -(g^{-1}dg) \wedge (g^{-1}dg) = -\Omega \wedge \Omega = -\frac{1}{2}[\Omega, \Omega]. \tag{7.166}$$

These are called the *Maurer-Cartan equations*.

If we restrict our attention to 0-form sections and orthonormal frames $e^u = \{e_1^U, e_2^U, \cdots e_n^U\}$ for $\mathbb{C}^N$ then $G = SU(N)$ and $\mathscr{G} = su(N)$ and hence the connection 1-forms $\{\omega_U\} \in su(N)$.

Since e^u is orthonormal

$$\begin{aligned}
\mathscr{L}_{e_k}\langle\langle e_i, e_j\rangle\rangle &= 0, \\
\langle\langle \nabla_{e_k} e_i, e_j\rangle\rangle &= \langle\langle \omega(e_k)e_i, e_j\rangle\rangle = \omega_i^j(e_k)^\dagger \\
\langle\langle e_i, \nabla_{e_k} e_j, \rangle\rangle &= \langle\langle e_i, \omega(e_k)e_j\rangle\rangle = \omega_j^i(e_k)
\end{aligned}$$

Since $\omega_i^j(e_k)^\dagger = -\omega_j^i(e_k)$ we have that

$$\mathscr{L}_{e_k}\langle\langle e_i, e_j\rangle\rangle = \langle\langle \nabla_{e_k} e_i, e_j\rangle\rangle + \langle\langle e_i, \nabla_{e_k} e_j, \rangle\rangle. \tag{7.167}$$

7.12 Connections on Principle Bundles

Consider the Principle bundle $\{P, M, \pi, G\}$ over the base manifold M. Then if $p = (q, g_U) \in \pi^{-1}(\mathscr{U})$ and $p = (q, g_V) \in \pi^{-1}(\mathscr{V})$ then $g_V = g_{VU} g_U$ where $g_{VU} \in G$.

Let $\{\omega_U\}$ be a locally defined $\mathscr{G}$-valued local 1-forms on M that satisfy the property

$$\omega_V = g_{UV}^{-1} \omega_U g_{UV} + g_{UV}^{-1} dg_{UV} \tag{7.168}$$

in overlaps. Here $g^{-1}\omega g = \mathrm{Ad}_{g^{-1}}\omega$.

Let $\{g_U\}$ be a section of P (that is $g_V = g_{VU} g_U$ in overlaps) and define

$$\nabla g_U \triangleq dg_U + \omega_U g_U. \tag{7.169}$$

Using (7.168) we can show

$$\begin{aligned}
\nabla g_U &= d(g_{UV} g_V) + (g_{UV}\omega_V g_{VU} + g_{UV} dg_{VU}) g_{UV} g_V \\
&= g_{UV}(dg_V + \omega_V g_V) + g_{UV}(g_{VU} dg_{UV} + dg_{VU} g_{UV}) g_V
\end{aligned}$$

$$= g_{UV}(dg_V + \omega_V g_V) = g_{UV}\nabla g_V.$$

Thus we see that (7.169) defines a connection on the principle bundle.

Let $q(t) \in M$ and g be a section of P then g is said to be covariant constant or parallel along $q(t)$ if

$$\nabla_{\dot{q}} g = \dot{g} + \omega(\dot{q})g = 0.$$

In the following we will show that the connection $\{\omega_U\}$ on P defines a global $\mathscr{G}$-valued 1-form and a 2-form on P so that it allows the TP to be split into a horizontal distribution and a vertical distribution.

Let $\pi : P \mapsto M$ be the projection map on P. In the patch $\pi^{-1}(\mathscr{U})$ let us define the $\mathscr{G}$ valued local 1-form on P (not on M as before) as follows

$$\omega_U^* = g_U^{-1}(\pi^* \omega_U) g_U + g_U^{-1} dg_U. \tag{7.170}$$

Since

$$\begin{aligned}
\omega_U^* &= g_U^{-1}(\pi^*(g_{UV}\omega_V g_{VU} + g_{UV} dg_{VU}))g_U + g_V^{-1} g_{UV}^{-1} d(g_{UV} g_V) \\
&= g_V^{-1}(\pi^*\omega_V)g_V + g_V^{-1} dg_{VU} g_U + g_V^{-1}(g_{UV}^{-1} dg_{UV} g_V + dg_V) \\
&= g_V^{-1}(\pi^*\omega_V)g_V - g_V^{-1} g_{VU}^{-1} dg_{UV} g_{VU} g_U + g_V^{-1} g_{UV}^{-1} dg_{UV} g_V + g_V^{-1} dg_V \\
&= g_V^{-1}(\pi^*\omega_V)g_V + g_V^{-1} dg_V = \omega_V^*
\end{aligned}$$

we see that the lifted $\mathscr{G}$ valued 1-forms $\{\omega_U^*\}$ are in fact the local representations of a global $\mathscr{G}$ valued 1-forms TP.

One can show that

$$\theta_U^* = g_U^{-1}\pi^*\theta_U g_U = g_V^{-1}(g_{UV}^{-1}\pi^*\theta_U g_{UV})g_V = g_V^{-1}\pi^*\theta_V g_V \tag{7.171}$$

$$= d\omega_U^* + \frac{1}{2}[\omega_U^*, \omega_U^*] \tag{7.172}$$

and hence that the lifted $\mathscr{G}$ valued 2-forms $\{\theta_U^*\}$ are in fact the local representations of a global $\mathscr{G}$ valued 2-form on TP.

Consider a patch $\mathscr{U} \times G = \pi^{-1}(\mathscr{U})$ and $p = (q, g_U) \in \mathscr{U} \times G$. Consider a vector field $\varphi(q, g_U)$ on P. In a local patch it takes the form $\varphi(q, g_U) = (X(q, g_U), g_U \xi_U(q, g_U))$ where $\xi_U(q, g_U) \in \mathscr{G}$. Then

$$\omega_U^*(\psi) = (g_U^{-1}\omega_U(X)g_U + \xi_U) \in \mathscr{G},$$

and thus that $g_U \omega_U^*(\psi) \in T_{g_U}G$ for any $\psi(q, g_U)$.

Let us define the distribution

$$\mathscr{H} = \{\varphi \in TP : \omega^*(\varphi) = 0\}.$$

We see that if $\varphi(q, g_U) = (0, g_U \xi_U)$ then $\omega^*(\varphi) \neq 0$ and hence that any $\varphi(q, g_U) = (0, g_U \xi_U) \notin \mathscr{H}$ and furthermore $g_U \omega^*(\varphi) = \varphi$. Thus that at every $p = (q, g_U) \in P$ the restriction of the distribution $\mathscr{H}_p$ is transversal to the fibre $T_{g_U}G$. Thus at every $p = (q, g_U) \in P$ we see that $T_p P \equiv \mathscr{H}_p \oplus T_{g_U}G$.

Define the right action $\phi : G \times P \mapsto P$ that is defined by $\phi_h(p) = \phi_h(q, g_U) \triangleq (q, g_U h)$. Then $T_{\phi_h(p)}P \equiv \mathscr{H}_{\phi_h(p)} \oplus T_{g_U h}G$. We will show below that $\mathscr{H}_{\phi_h(p)} = T_p\phi_h(\mathscr{H}_p)$. This follows from the following calculation:

$$i_{(\dot{q}, \dot{g}_U h)}\omega^*_{(q, g_U h)} = h^{-1}g_U^{-1}\omega_U(\dot{q})g_U h + h^{-1}g_U^{-1}\dot{g}_U h = \mathrm{Ad}_{h^{-1}}\left(i_{(\dot{q}, \dot{g}_U)}\omega^*_{(q, g_U)}\right).$$

In compact notation this means

$$\phi_h^*\omega^* = \mathrm{Ad}_{h^{-1}}\omega^*. \tag{7.173}$$

This says that if $\dot{p}(t) \in \mathscr{H}$ then so is $T_{p(t)}\phi_h\dot{p}(t) \in \mathscr{H}$. This says that if $p(t)$ is a horizontal curve then so is $\phi_h(p(t))$ for any $h \in G$.

Let $p(t) = (q(t), g_U(t))$ be a curve on P. The tangent vector to $p(t)$ is then $\dot{p} = (\dot{q}, \dot{g})$. If $p(t)$ is such that $\dot{p} = (\dot{q}, \dot{g}_U) \in \mathscr{H}$ then we have that

$$\omega_U^*(\dot{p}) = (g_U^{-1}\pi^*\omega_U(\dot{q})g_U + g_U^{-1}\dot{g}_U) = 0,$$

and hence along $p(t)$

$$\nabla_{\dot{q}}g_U = \dot{g}_U + \omega_U(\dot{q})g_U = 0.$$

Let X be a vector field on M. Then ψ_X is called the *horizontal lift* of X if $\psi_X \in \mathscr{H}$ and $\pi_*\psi_X = X$. Thus necessarily the horizontally lifted field of X must take the form,

$$\psi_X = (X, -\omega_U(X)g_U).$$

Consider a closed path $\gamma : [0, T] \mapsto M$ then the parallel transport of g_0 along $\gamma(t)$ starting from $\gamma(0)$ is the solution to $\nabla_{\dot{\gamma}}g = 0$ with initial condition $g(0) = g_0$. If $\mathscr{H}$ is integrable and $\gamma(t)$ is a sufficiently small closed curve then $g(T) = g_0$. On the other hand if $\mathscr{H}$ is not integrable then it may turn out that $g(T) \neq g_0$.

It can be shown that $\mathscr{H}$ is integrable if and only of the curvature $\theta_U^* = 0$.

7.13 Connection on Associated Bundles

Consider a vector bundle $\{E, M, \pi, \mathbb{C}^N\}$ with a structure group G. Let $\rho : G \to GL(N)$ be some representation of G where $GL(N)$ is the space of $N \times N$ invertible matrices that act on $\mathbb{C}^N$ by linear transformations. We may then define a new vector bundle $\{E_\rho, M, \pi, \mathbb{C}^N\}$ such that in an overlap $\mathscr{U} \cap \mathscr{V}$ with transition function $c_{UV} \in G$ of the vector bundle the representations of a r-form section of the new associated bundle are related by

$$\psi_V = \rho(c_{VU})\psi_U.$$

In the patch $\pi^{-1}(\mathscr{U})$ let ω_U be a locally defined $\mathscr{G}$ valued connection 1-form of the vector bundle E and θ_U be the corresponding locally defined curvature 2-form. We have seen that in an overlap $\mathscr{U} \cap \mathscr{V}$ these local forms are related by

$$\omega_V = c_{UV}^{-1}\omega_U c_{UV} + c_{UV}^{-1} dc_{UV},$$
$$\theta_V = c_{UV}^{-1}\theta_U c_{UV}.$$

Thus we see that the the connection 1-forms $\{\omega_U\}$ nor the curvature 2-forms $\{\theta_U\}$ correspond to global forms of M nor are they form sections of the bundle E. However, notice that

$$\theta_V = c_{UV}^{-1}\theta_U c_{UV} = \rho(c_{VU})\theta_U$$

where $\rho = \mathrm{Ad}$ and hence that the curvature 2-forms $\{\theta_U\}$ are the local representations of a global 2-form section of the Ad-bundle associated with the vector bundle E.

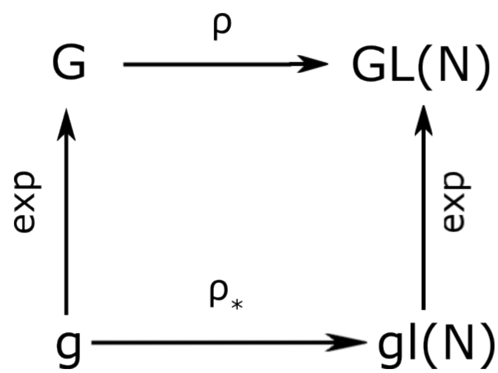

Fig. 7.1 Associated Bundles

Define

$$\Omega_U = \rho_*(\omega_U) \triangleq T_e\rho\cdot\omega_U = \left.\frac{d}{dt}\right|_{t=0}\rho(\exp(t\zeta)), \tag{7.174}$$

where $\zeta \in \mathscr{G}$. We see that Ω_U is a local Lie algebra valued 1-form of the associated bundle E_ρ. We see that

$$\begin{aligned}\Omega_V &= \rho_*(\omega_V) = \rho_*(c_{UV}^{-1}\omega_U c_{UV} + c_{UV}^{-1}dc_{UV}) = \rho(c_{UV}^{-1})\rho_*(\omega_U)\rho(c_{UV}) + \rho_*(c_{UV}^{-1}dc_{UV})\\ &= \rho(c_{UV}^{-1})\rho_*(\omega_U)\rho(c_{UV}) + \rho(c_{UV}^{-1})d\rho(c_{UV}) = \rho(c_{UV})^{-1}\Omega_U\rho(c_{UV}) + \rho(c_{UV})^{-1}d\rho(c_{UV})\end{aligned}$$

and hence that the collection $\{\Omega_U\}$ defined by (7.174) is a connection 1-form of the associated bundle E_ρ.

Then for a given r-form section of the associated bundle E_ρ if we define

$$\nabla\psi_U = d\psi_U + \Omega_U\wedge\psi_U, \tag{7.175}$$

then since

$$\begin{aligned}\nabla\psi_V &= d\psi_V + \Omega_V\wedge\psi_V\\ &= d(\rho(c_{VU})\psi_U) + \rho(c_{UV})^{-1}\Omega_U\rho(c_{UV})\wedge\rho(c_{VU})\psi_U + \rho(c_{UV})^{-1}d\rho(c_{UV})\wedge\rho(c_{VU})\psi_U\\ &= d(\rho(c_{VU})\psi_U) + \rho(c_{VU})\Omega_U\wedge\psi_U + \rho(c_{UV})^{-1}d\rho(c_{UV})\wedge\rho(c_{VU})\psi_U\end{aligned}$$

$$= \rho(c_{VU})(d\psi_U + \Omega_U \wedge \psi_U) = \rho(c_{VU})\nabla\psi_V$$

Thus we see that (7.175) defines a covariant derivative on sections of the associated bundle E_ρ.

Consider the case where ρ is $\mathrm{Ad}: G \to GL(\mathscr{G})$. Then we know that $\mathrm{Ad}_* : \mathscr{G} \to \mathscr{G}$ is given by

$$\mathrm{Ad}_*(\zeta) = \mathrm{ad}_\zeta = [\zeta, \cdot].$$

Hence we have that the connection on the associated Ad-bundle is given by $\Omega = \mathrm{Ad}_*(\omega) = [\omega, \cdot]$ and hence the covariant derivative of a Lie algebra valued section ψ of the Ad-bundle is given by

$$\nabla\psi = d\psi + [\omega, \psi]. \tag{7.176}$$

A Lie algebra valued section of the Ad-bundle will be called a Lie algebra valued 0-form section of the type $\mathrm{Ad}(G)$. When ψ is a 0-form section (a plain section) then

$$\nabla_X\psi = d\psi(X) + [\omega(X), \psi]. \tag{7.177}$$

We have seen that the curvature

$$\theta = d\omega + \omega\wedge\omega = d\omega + \frac{1}{2}[\omega, \omega]$$

is a globally defined 2-form section of the Ad-bundle. Thus we find that the covariant derivative of the curvature is

$$\begin{aligned}\nabla\theta &= d\theta + [\omega, \theta] = d\omega\wedge\omega - \omega\wedge d\omega + [\omega, d\omega] + [\omega, \omega\wedge\omega]\\ &= d\omega\wedge\omega - \omega\wedge d\omega + (\omega\wedge d\omega - d\omega\wedge\omega) + (\omega\wedge\omega\wedge\omega - \omega\wedge\omega\wedge\omega)\\ &= 0.\end{aligned}$$

This is known as the *Bianchi identity*.

Notice that $\{E_{\mathrm{Ad}}, M, \pi, \mathscr{G}\}$ is a bundle with transition functions $\mathrm{Ad}_{g_{UV}}$ and $g_{UV}(g) = g$. Thus we now have a covariant derivative for $\mathscr{G}$-valued r-form sections, ψ, on M that is given by

$$\nabla\psi = d\psi + [\omega, \psi]. \tag{7.178}$$

For instance consider the bundle $\{E_{\mathrm{Ad}}, SO(3), \pi, so(3)\}$ that is associated with the vector bundle $\{TSO(3), SO(3), \pi, so(3)\}$. We identify $(R, \zeta^L) \in G\times so(3)$ with $(R, \zeta^R) \in G\times so(3)$ if $\zeta^R = \mathrm{Ad}_R\zeta^L$. Thus if $\omega = \{\omega^L, \omega^R\}$ are locally defined $so(3)$ valued 1-forms such that

$$\omega^R = R\omega^L R^T + RdR^T,$$

Then they represent a connection on the bundle E. Hence we have that $\mathrm{Ad}_*\omega = \{\mathrm{Ad}_*\omega^L, \mathrm{Ad}_*\omega^R\} = \{\mathrm{ad}_{\omega^L}, \mathrm{ad}_{\omega^R}\}$ is the associated connection 1-form on E_{Ad}. Thus we have that

$$\nabla\eta^a = d\eta^a + [\omega^a(\cdot), \eta^a].$$

The curvature forms $\theta = \{\theta^L, \theta^R\}$ satisfy

$$\theta^R = R\theta^L R^T.$$

7.13.1 Flat Connection

Consider the special case of $\omega^R = 0$ then

$$\omega^L = -dR^T R = R^T dR.$$

Thus we have

$$\begin{aligned}\nabla_{\zeta^R}\eta^R &= d\eta^R,\\ \nabla_{\zeta^L}\eta^L &= d\eta^L + [\zeta^L, \eta^L].\end{aligned}$$

Since $\omega^R = 0$ we see that $\theta^R = 0$. Hence $\theta^L = 0$.

7.13.2 The Levi-Civita Connection

Let $\mathbb{I}$ be a left-invariant metric on G and consider the frame $e = \{e_1, e_2, \cdots, e_n\}$ that is orthonormal with respect to $\mathbb{I}$ for $\mathscr{G}$. Then since

$$\nabla_\zeta \eta = d\eta(\zeta) \pm \frac{1}{2}[\zeta, \eta],$$

we see that the 1-forms $\omega = \{\frac{1}{2}\Omega^L, -\frac{1}{2}\Omega^R\}$ correspond to the Levi-Civita connection. We then have that $\theta = \{\theta^L, \theta^R\} = \{\frac{1}{8}[\Omega^L, \Omega^L], \frac{3}{8}[\Omega^R, \Omega^R]\}$

7.13.3 Connections and Nonholonomic Mechanical Systems

Consider a mechanical system with configuration space $\mathscr{P}$ equal to a principle bundle $\{P, M, \pi, G\}$ such that its trajectories $(p(t), \dot{p}(t))$ are constrained to satisfy the condition $\dot{p} \in \mathscr{H}$. Note that $p(t) = (q(t), g(t))$ and that $\dot{p} \in \mathscr{H}$ implies

$$\dot{g} = -\omega(\dot{q})g = -g(\mathrm{Ad}_{g-1}\omega(\dot{q})).$$

These are the nonholonomic constraints. Let ω_L be the matrix of 1-forms corresponding to the Levi-Civita connection on $\mathscr{P}$. Then we have

$$\nabla_{\dot{p}}\dot{p} = \mathbb{I}^{-1}(f^e + \Lambda\omega^*) = \mathbb{I}^{-1}\left(f^e + \Lambda(g^{-1}\omega g + g^{-1}dg)\right)$$

and thus the equations of motion are

$$\omega^*(\nabla_{\dot{p}}\dot{p}) = \omega^*(\mathbb{I}^{-1}f^e) \tag{7.179}$$

$$\dot{g} = -\omega(\dot{q})g \tag{7.180}$$

Rolling Ball

The configuration space $\mathcal{Q} = \mathbb{R}^3 \times SO(3)$. A particular configuration $q = (o, R)$ and $\zeta = q^{-1}\dot{q} = (V, \Omega)$.

$$\text{KE} = \frac{M}{2}||V||^2 + \frac{1}{2}\langle \mathbb{I}\Omega, \Omega\rangle = \frac{M}{2}||V||^2 + \frac{1}{2}\Omega^T \mathbb{I}\Omega$$

Constraints

$$\begin{aligned} V &= -R^T \widehat{e}_3 R\Omega. \\ \omega^* &= \sigma_v + (R^T \widehat{e}_3 R)\sigma_\Omega \end{aligned}$$

The connection

$$\nabla_\zeta \zeta = (\dot{V} - V \times \Omega\ ,\ \dot{\Omega} - \mathbb{I}^{-1}(\mathbb{I}\Omega \times \Omega)).$$

The forces

$$\text{Generalized \ forces} = F = f\sigma_V + \tau\sigma_\Omega.$$

$$\omega^*(\nabla_\zeta \zeta) = \dot{V} - V \times \Omega + (R^T \widehat{e}_3 R)(\dot{\Omega} - \mathbb{I}^{-1}(\mathbb{I}\Omega \times \Omega))$$

Equations of motion

$$\begin{aligned} \dot{R} &= R\widehat{\Omega}, \\ \dot{o} &= RV, \\ V &= -R^T \widehat{e}_3 R\Omega, \\ \dot{V} - V \times \Omega + (R^T \widehat{e}_3 R)(\dot{\Omega} - \mathbb{I}^{-1}(\mathbb{I}\Omega \times \Omega)) &= \omega^*(\mathbb{I}_G^{-1} F) \end{aligned}$$

$$\begin{aligned} \dot{V} &= \widehat{\Omega} R^T \widehat{e}_3 R\Omega - R^T \widehat{e}_3 R\widehat{\Omega}\Omega - R^T \widehat{e}_3 R\dot{\Omega} \\ &= \widehat{\Omega} R^T \widehat{e}_3 R\Omega - R^T \widehat{e}_3 R\dot{\Omega} \\ &= \Omega \times (R^T(e_3 \times R\Omega)) - R^T \widehat{e}_3 R\dot{\Omega} = \Omega \times (R^T e_3 \times \Omega) - R^T \widehat{e}_3 R\dot{\Omega} \\ &= -\widehat{\Omega}^2 R^T e_3 - R^T \widehat{e}_3 R\dot{\Omega} \end{aligned}$$

$$V \times \Omega = -R^T \widehat{e}_3 R\Omega \times \Omega = -R^T(e_3 \times (R\Omega \times \Omega)),$$

Hence eom

$$\begin{aligned} \dot{R} &= R\widehat{\Omega}, \\ \dot{o} &= -\widehat{e}_3 R\Omega, \\ V &= -R^T \widehat{e}_3 R\Omega, \\ -\widehat{\Omega}^2 R^T e_3 - R^T \widehat{e}_3 R\dot{\Omega} + R^T(e_3 \times (R\Omega \times \Omega)) + (R^T \widehat{e}_3 R)(\dot{\Omega} - \mathbb{I}^{-1}(\mathbb{I}\Omega \times \Omega)) &= \omega^*(\mathbb{I}_G^{-1} F) \end{aligned}$$

$$\dot{R} = R\widehat{\Omega},$$

$$\begin{aligned}
\dot{o} &= -\widehat{e}_3 R\Omega, \\
\omega^*(\mathbb{I}_G^{-1}F) &= -\widehat{\Omega}^2 R^T e_3 + R^T(e_3 \times (R\Omega \times \Omega)) - (R^T\widehat{e}_3 R)(\mathbb{I}^{-1}(\mathbb{I}\Omega \times \Omega)) \\
V &= -R^T\widehat{e}_3 R\Omega,
\end{aligned}$$

$$\delta g = -\omega(\delta q)g = -g(\mathrm{Ad}_{g-1}\omega(\delta q)).$$

$$\begin{aligned}
\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q}\right)\delta q + \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{g}}\right) - \frac{\partial L}{\partial g}\right)\delta g = 0 \\
\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q}\right)\delta q - \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{g}}\right) - \frac{\partial L}{\partial g}\right)\omega(\delta q)g = f_q\delta q + f_g\delta g = f_q\delta q - f_g\omega(\delta q)g \\
\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} - f_q\right)\delta q - \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{g}}\right) - \frac{\partial L}{\partial g} - f_g\right)\omega(\delta q)g = 0.
\end{aligned}$$

If the group is abelian $\omega(\dot{q}) = A\dot{q}$ and

$$\begin{aligned}
\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} - f_q\right)\delta q - \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{g}}\right) - \frac{\partial L}{\partial g} - f_g\right)A\delta q = 0 \\
\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} - f_q = \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{g}}\right) - \frac{\partial L}{\partial g} - f_g\right)A
\end{aligned}$$

Thus equations of motion are

$$\begin{aligned}
\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} &= \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{g}}\right) - \frac{\partial L}{\partial g}\right)A + f_q + f_g A \\
\dot{g} &= -A\dot{q}
\end{aligned}$$

On the other hand we see that if ω_U corresponds to the Levi-Civita connection on M and $\omega_U \in \mathscr{G}$ (that is we use orthonormal frames for T_mM) then

$$\nabla_{\dot{q}}\dot{q} = \ddot{q} + \omega_U(\dot{q})\dot{q}.$$

7.14 Affine Connections on Lie Groups

We know that TG is an Ad-bundle with two patches corresponding to left and right translations. Consider a section $\{\zeta^L, \zeta^R\}$ of the bundle (that is $\zeta^R = \mathrm{Ad}_g\zeta^L$).

$$\nabla_X Y = g \cdot \nabla^L_{\zeta^L}\eta^L = \nabla^R_{\zeta^R}\eta^R \cdot g$$

Define the isomorphisms $\mathbb{I}_L : \mathscr{G} \mapsto \mathscr{G}^*$ and $\mathbb{I}_R : \mathscr{G} \mapsto \mathscr{G}^*$ as follows:

$$\langle\langle \dot{g}, \dot{h} \rangle\rangle = \langle \mathbb{I}_L \zeta^L, \eta^L \rangle = \langle \mathbb{I}_R \zeta^R, \eta^R \rangle$$

where $\dot{g} = g \cdot \zeta^L = \zeta^R \cdot g$ and $\dot{h} = g \cdot \eta^L = \eta^R \cdot g$. Since

$$\langle \mathbb{I}_L \zeta^L, \eta^L \rangle = \langle \mathbb{I}_L \mathrm{Ad}_{g^{-1}} \zeta^R, \mathrm{Ad}_{g^{-1}} \eta^R \rangle = \langle \mathrm{Ad}^*_{g^{-1}} \mathbb{I}_L \mathrm{Ad}_{g^{-1}} \zeta^R, \eta^R \rangle,$$

we see that

$$\mathbb{I}_R = \mathrm{Ad}^*_{g^{-1}} \mathbb{I}_L \mathrm{Ad}_{g^{-1}}.$$

Let $\mathbf{E}^L = g \cdot \mathbf{e}$ and $\mathbf{E}^R = \mathbf{e} \cdot g$.

$$\nabla Y = E^L(d\zeta_Y^L + \omega^L(\cdot)\zeta_Y^L) = E^R(d\zeta_Y^R + \omega^R(\cdot)\zeta_Y^R)$$

where

$$\omega : T_g G \mapsto gl(\mathscr{G}).$$

and since

$$\begin{aligned}
\mathrm{Ad}_g(d\zeta_Y^L + \omega^L(\cdot)\zeta_Y^L) &= (d\zeta_Y^R + \omega^R(\cdot)\zeta_Y^R) \\
&= d(\mathrm{Ad}_g \zeta_Y^L) + \omega^R(\cdot)\mathrm{Ad}_g \zeta_Y^L = \mathrm{Ad}_g d\zeta_Y^L + d\mathrm{Ad}_g \zeta_Y^L + \omega^R(\cdot)\mathrm{Ad}_g \zeta_Y^L \\
&= \mathrm{Ad}_g \left(d\zeta_Y^L + \mathrm{Ad}_{g^{-1}} d\mathrm{Ad}_g \zeta_Y^L + \mathrm{Ad}_{g^{-1}} \omega^R(\cdot)\mathrm{Ad}_g \zeta_Y^L \right)
\end{aligned}$$

we see that

$$\omega^R = \mathrm{Ad}_g d\mathrm{Ad}_{g^{-1}} + \mathrm{Ad}_g \omega^L(\cdot)\mathrm{Ad}_{g^{-1}}.$$

Hence

$$\begin{aligned}
\omega^R(\dot{g}) &= \mathrm{Ad}_g \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{g^{-1}\exp(-t\eta_Y^R)} + \mathrm{Ad}_g \omega^L(\dot{g})\mathrm{Ad}_{g^{-1}} \\
&= -\mathrm{ad}_{\eta_Y^R} + \mathrm{Ad}_g \omega^L(\dot{g})\mathrm{Ad}_{g^{-1}}.
\end{aligned}$$

7.15 The Riemannian Structure On Lie Groups

If the Lie group G is equipped with the structure of a Riemannian manifold $(G, << \cdot, \cdot >>)$. Let $\{e_i\}$ be any basis for the Lie algebra $\mathscr{G}$ and let $\{E_i^L(g) = g \cdot e_i\}$ be the associated left invariant basis vector field on G and $\{E_i^R(g) = e_i \cdot g\}$ be the associated right invariant basis vector field on G. The left invariant 1-form field dual to $\{E_i\}$ will be denoted by $\{\sigma^i\}$ (that is $\sigma^i(E_j) = \delta_j^i$). Now $[e_i, e_j]_{\mathscr{G}} = C_{ij}^k e_k$, where C_{ij}^k are the structure constants of the Lie algebra $\mathscr{G}$ (note that $C_{ij}^k = -C_{ji}^k$), then $[E_i^L, E_j^L] = C_{ij}^k E_k^L$. In what follows we will specialize the notions of a Levi-Civita connection, Riemannian curvature and the notion of a distance function to Lie groups equipped with a left invariant metric. For any vectors $v, w \in T_g G$ at $g \in G$ and locally defined smooth vector field $X = \xi^k(g)E_k^L \in \mathscr{X}(G)$ where $\xi^k(g) \in C^\infty(G)$ and a smooth function $f \in C^\infty(G)$ the operator $\nabla : T_gG \times \mathscr{X}(G) \to T_gG$ that satisfies the following properties

$$\nabla_{v+w}X = \nabla_v X + \nabla_w X, \tag{7.181}$$
$$\nabla_v fX = (\mathscr{L}_v f)X + f\nabla_v X, \tag{7.182}$$

is defined to be an affine connection on G.

On any Riemannian manifold it is known that there exists a unique connection that is metric and torsion free [**?**, **?**, **?**, **?**]. That is, for the locally defined smooth vector fields X,Y,Z on G the following conditions are satisfied,

$$\mathscr{L}_Z << X,Y >> = << \nabla_Z X, Y >> + << X, \nabla_Z Y >>, \tag{7.183}$$
$$[X,Y] = \nabla_X Y - \nabla_Y X, \tag{7.184}$$

where $\mathscr{L}$ represents the Lie derivative. This connection is referred to as the Riemannian or Levi-Civita connection. Condition (7.183) says that the connection is metric and the condition (7.184) says that the metric is torsion free. Combining these two conditions, the Levi-Civita connection is uniquely given by the Koszul formula [**?**],

$$\begin{aligned} 2 << \nabla_X Y, Z >> = \mathscr{L}_X << Y,Z >> + \mathscr{L}_Y << Z,X >> - \mathscr{L}_Z << X,Y >> \\ + << [X,Y],Z >> - << [Y,Z],X >> + << [Z,X],Y >> . \end{aligned} \tag{7.185}$$

7.15.1 The Levi-Civita connection of a left invariant metric

We first compute the connection on left invariant vector fields on G. Then (7.185) yields,

$$\begin{aligned} 2 << \nabla_{X^L_\zeta} X^L_\eta, X^L_\xi >> &= \mathscr{L}_{X^L_\zeta} << X^L_\eta, X^L_\xi >> + \mathscr{L}_{X^L_\eta} << X^L_\xi, X^L_\zeta >> - \mathscr{L}_{X^L_\xi} << X^L_\zeta, X^L_\eta >> \\ &\quad + << [X^L_\zeta, X^L_\eta], X^L_\xi >> - << [X^L_\eta, X^L_\xi], X^L_\zeta >> + << [X^L_\xi, X^L_\zeta], X^L_\eta >>, \\ &= \mathscr{L}_{X^L_\zeta} < \mathbb{I}\eta, \xi > + \mathscr{L}_{X^L_\eta} < \mathbb{I}\zeta, \xi > - \mathscr{L}_{X^L_\xi} < \mathbb{I}\zeta, \eta > \\ &\quad + < \mathbb{I}[\zeta,\eta]_{\mathscr{G}}, \xi > - < \mathbb{I}\zeta, [\eta,\xi]_{\mathscr{G}} > + < \mathbb{I}\eta, [\xi,\zeta]_{\mathscr{G}} >, \\ &= < \mathbb{I}[\zeta,\eta]_{\mathscr{G}}, \xi > - < \mathbb{I}\zeta, [\eta,\xi]_{\mathscr{G}} > + < \mathbb{I}\eta, [\xi,\zeta]_{\mathscr{G}} >, \\ &= < \mathbb{I}[\zeta,\eta]_{\mathscr{G}}, \xi > - < ad^*_\eta \mathbb{I}\zeta, \xi > - < ad^*_\zeta \mathbb{I}\eta, \xi >, \end{aligned}$$

Note that $<< \nabla_{X^L_\zeta} X^L_\eta, X^L_\xi >> = < \mathbb{I}\,(g^{-1} \cdot \nabla_{X^L_\zeta} X^L_\eta), \xi >$. Hence, the Levi-Civita connection of a left invariant metric on left invariant vector fields on G is given by $\nabla_{X^L_\zeta} X^L_\eta = g \cdot \nabla_\zeta \eta$ where

$$\nabla_\zeta \eta = \frac{1}{2}\Big\{[\zeta,\eta]_{\mathscr{G}} - \mathbb{I}^{-1}(ad^*_\zeta \mathbb{I}\eta + ad^*_\eta \mathbb{I}\zeta)\Big\}. \tag{7.186}$$

Using this expression and $[e_i, e_j]_{\mathscr{G}} = C^k_{ij} e_k$, where C^k_{ij} are the structure constants of the Lie algebra $\mathscr{G}$, we have that,

$$\nabla_{e_i} e_j = \frac{1}{2}\left\{[e_i, e_j]_{\mathscr{G}} - \mathbb{I}^{-1}(ad^*_{e_i}\mathbb{I}e_j + ad^*_{e_j}\mathbb{I}e_i)\right\},$$
$$= \frac{1}{2}\left(C^k_{ij} - \mathbb{I}^{ks}(\mathbb{I}_{ir}C^r_{js} + \mathbb{I}_{jr}C^r_{is})\right)e_k \triangleq \omega^k_{ij}e_k, \tag{7.187}$$

where the ω^k_{ij} given by

$$\omega^k_{ij} = \frac{1}{2}\left(C^k_{ij} - \mathbb{I}^{ks}(\mathbb{I}_{ir}C^r_{js} + \mathbb{I}_{jr}C^r_{is})\right). \tag{7.188}$$

are defined to be the the connection coefficients. For any vectors v, w at $g \in G$ and locally defined smooth vector field $X = \xi^k(g)E^L_k$ where $\xi^k(g) \in C^\infty(G)$ and a smooth function $f \in C^\infty(G)$ the conditions (7.181)–(7.182) defines the Levi-Civita or Riemannian connection on G as,

$$\nabla_v X = d\xi^k(v)E^L_k + v^i\xi^j\nabla_{E^L_i}E^L_j = d\xi^k(v)E^L_k + v^i\xi^j g\cdot\nabla_{e_i}e_j = d\xi^k(v)E^L_k + g\cdot\nabla_{v^ie_i}\xi^ie_j. \tag{7.189}$$

Note that $\eta = g^{-1}\cdot v = g^{-1}\cdot v^iE_i = g^{-1}\cdot v^ig\cdot e_i = v^ie_i \in \mathscr{G}$ and similarly that $\xi(g) = g^{-1}\cdot X \in \mathscr{G}$. Thus we have from the above expression that

$$\nabla_v X = g\cdot\left(d\xi^k(v)e^k + \frac{1}{2}\left([\eta, \xi(g)]_{\mathscr{G}} - \mathbb{I}^{-1}\left(ad^*_\eta\mathbb{I}\xi(g) + ad^*_{\xi(g)}\mathbb{I}\eta\right)\right)\right). \tag{7.190}$$

The Jacobi Lie-bracket of two vector fields $X = \xi^k(g)E_k$, $Y = \eta^k(g)E_k$ on G will also be given by (7.184) and (7.190) as follows,

$$[X, Y] = \left\{d\eta^k(X) - d\xi^k(Y)\right\}E_k + g\cdot[\xi(g), \eta(g)]_{\mathscr{G}}. \tag{7.191}$$

We also see from (7.187) that

$$\nabla e_j = \frac{1}{2}\left\{[\cdot, e_j]_{\mathscr{G}} - \mathbb{I}^{-1}(ad^*_{(\cdot)}\mathbb{I}e_j + ad^*_{e_j}\mathbb{I}(\cdot))\right\} = e_k\omega^k_{ij}\sigma^i = e_k\omega^k_j(\cdot),$$

Hence we can define

$$\nabla^L e = e\omega^L,$$

where

$$\omega^L : \mathscr{G} \mapsto gl(\mathscr{G}) \Longrightarrow \omega^k_j = \sum_{i=1}^N \omega^k_{ij}\sigma^i.$$

Here ω^L is the matrix of connection 1-forms that correspond to the Levi-Civita connection of a left invariant metric. Then $\nabla^L e\zeta^L = e(d\zeta^L + \omega^L(\cdot)\zeta^L)$ and we may define

$$\nabla^L\zeta^L = d\zeta^L + \omega^L(\cdot)\zeta^L.$$

and hence

$$\omega^L(\cdot)\eta^L = \frac{1}{2}\left\{[\cdot, \eta^L]_{\mathscr{G}} - \mathbb{I}_L^{-1}(ad^*_{(\cdot)}\mathbb{I}_L\eta^L + ad^*_{\eta^L}\mathbb{I}_L(\cdot))\right\}.$$

Since for $\dot{g} = \zeta^R \cdot g = g \cdot \zeta^L$

$$\omega^R(\dot{g}) = -\mathrm{ad}_{\zeta^R} + \mathrm{Ad}_g \omega^L(\zeta^L)\mathrm{Ad}_{g^{-1}}.$$

we have that

$$\begin{aligned}\omega^R(\zeta^R)\eta^R &= -\mathrm{ad}_{\zeta^R}\eta^R + \mathrm{Ad}_g \omega^L(\zeta^L)\mathrm{Ad}_{g^{-1}}\eta^R \\ &= -\mathrm{ad}_{\zeta^R}\eta^R + \mathrm{Ad}_g \omega^L(\zeta^L)\eta^L.\end{aligned}$$

Let us explicitly compute $\mathrm{Ad}_g \omega^L(\zeta^L)\eta^L$.

$$\mathrm{Ad}_g \omega^L(\zeta^L)\eta^L = \frac{1}{2}\left\{\mathrm{Ad}_g \mathrm{ad}_{\zeta^L}\eta^L - \mathrm{Ad}_g \mathbb{I}_L^{-1}\left(ad^*_{\zeta^L}\mathbb{I}_L\eta^L + ad^*_{\eta^L}\mathbb{I}_L\zeta^L\right)\right\}.$$

Since

$$\mathrm{Ad}_g \mathrm{ad}_{\zeta^L}\eta^L = \mathrm{ad}_{\mathrm{Ad}_g\zeta^L}\mathrm{Ad}_g\eta^L = \mathrm{ad}_{\zeta^R}\eta^R$$

and

$$\begin{aligned}\langle \mathrm{ad}^*_{\zeta^L}\mathbb{I}_L\eta^L, \nu^L\rangle &= \langle \mathrm{ad}^*_{\mathrm{Ad}_{g^{-1}}\zeta^R}\mathbb{I}_L\mathrm{Ad}_{g^{-1}}\eta^R, \mathrm{Ad}_{g^{-1}}\nu^R\rangle = \langle \mathbb{I}_L\mathrm{Ad}_{g^{-1}}\eta^R, \mathrm{ad}_{\mathrm{Ad}_{g^{-1}}\zeta^R}\mathrm{Ad}_{g^{-1}}\nu^R\rangle \\ &= \langle \mathbb{I}_L\mathrm{Ad}_{g^{-1}}\eta^R, \mathrm{Ad}_{g^{-1}}\mathrm{ad}_{\zeta^R}\nu^R\rangle = \langle \mathrm{Ad}^*_{g^{-1}}\mathbb{I}_L\mathrm{Ad}_{g^{-1}}\eta^R, \mathrm{ad}_{\zeta^R}\nu^R\rangle \\ &= \langle \mathrm{ad}^*_{\zeta^R}\mathbb{I}_R\eta^R, \nu^R\rangle = \langle \mathrm{ad}^*_{\zeta^R}\mathbb{I}_R\eta^R, \mathrm{Ad}_g\nu^L\rangle \\ &= \langle \mathrm{Ad}^*_g\mathrm{ad}^*_{\zeta^R}\mathbb{I}_R\eta^R, \nu^L\rangle\end{aligned}$$

and hence

$$\mathrm{ad}^*_{\zeta^L}\mathbb{I}_L\eta^L = \mathrm{Ad}^*_g\mathrm{ad}^*_{\zeta^R}\mathbb{I}_R\eta^R.$$

Thus we have

$$\begin{aligned}\mathrm{Ad}_g \omega^L(\zeta^L)\eta^L &= \frac{1}{2}\left\{\mathrm{ad}_{\zeta^R}\eta^R - \mathrm{Ad}_g\mathbb{I}_L^{-1}\mathrm{Ad}^*_g\left(ad^*_{\zeta^R}\mathbb{I}_R\eta^R + ad^*_{\eta^R}\mathbb{I}_R\zeta^R\right)\right\} \\ &= \frac{1}{2}\left\{\mathrm{ad}_{\zeta^R}\eta^R - \mathbb{I}_R^{-1}\left(ad^*_{\zeta^R}\mathbb{I}_R\eta^R + ad^*_{\eta^R}\mathbb{I}_R\zeta^R\right)\right\}.\end{aligned}$$

Thus we finally have

$$\omega^R(\zeta^R)\eta^R = -\frac{1}{2}\left\{\mathrm{ad}_{\zeta^R}\eta^R + \mathbb{I}_R^{-1}\left(ad^*_{\zeta^R}\mathbb{I}_R\eta^R + ad^*_{\eta^R}\mathbb{I}_R\zeta^R\right)\right\},$$

and hence that

$$\nabla_\xi \eta \triangleq d\eta(\xi) + \frac{1}{2}\left(\pm\,\mathrm{ad}_\xi\eta - \mathbb{I}^{-1}\left(\mathrm{ad}^*_\xi\mathbb{I}\eta + \mathrm{ad}^*_\eta\mathbb{I}\xi\right)\right),$$

where the $+$ results in $\nabla^L_{\xi^L}\eta^L$ and the $-$ results in $\nabla^R_{\xi^R}\eta^R$.

7.15.2 The Levi-Civita connection of a right Invariant metric

Similar to how we proceeded in the previous section we will first compute the connection on right invariant vector fields on G. Then as we have seen in the previous section will allow us to compute the connection on general vectorfields.

The Kozul formula (7.185) yields,

$$\begin{aligned}
2 << \nabla_{X^R_\zeta} X^R_\eta \,, X^R_\xi >> &= \mathscr{L}_{X^R_\zeta} << X^R_\eta, X^R_\xi >> + \mathscr{L}_{X^R_\eta} << X^R_\xi, X^R_\zeta >> - \mathscr{L}_{X^R_\xi} << X^R_\zeta, X^R_\eta >> \\
&\quad + << [X^R_\zeta, X^R_\eta] \,, X^R_\xi >> - << [X^R_\eta, X^R_\xi] \,, X^R_\zeta >> + << [X^R_\xi, X^R_\zeta] \,, X^R_\eta >>, \\
&= \mathscr{L}_{X^R_\zeta} < \mathbb{I}\eta, \xi > + \mathscr{L}_{X^R_\eta} < \mathbb{I}\zeta, \xi > - \mathscr{L}_{X^R_\xi} < \mathbb{I}\zeta, \eta > \\
&\quad - < \mathbb{I}[\zeta,\eta]_{\mathscr{G}} \,, \xi > + < \mathbb{I}\zeta \,, [\eta,\xi]_{\mathscr{G}} > - < \mathbb{I}\eta \,, [\xi,\zeta]_{\mathscr{G}} >, \\
&= - < \mathbb{I}[\zeta,\eta]_{\mathscr{G}} \,, \xi > + < \mathbb{I}\zeta \,, [\eta,\xi]_{\mathscr{G}} > - < \mathbb{I}\eta \,, [\xi,\zeta]_{\mathscr{G}} >, \\
&= - < \mathbb{I}[\zeta,\eta]_{\mathscr{G}} \,, \xi > + < ad^*_\eta \mathbb{I}\zeta, \xi > + < ad^*_\zeta \mathbb{I}\eta, \xi >,
\end{aligned}$$

Note that $<< \nabla_{X^R_\zeta} X^R_\eta \,, X^R_\xi >> = < \mathbb{I}\,(\nabla_{X^R_\zeta} X^R_\eta) \cdot g^{-1} \,, \xi >$. Hence, the Levi-Civita connection of a right invariant metric acting on right invariant vector fields on G is given by $\nabla_{X^R_\zeta} X^R_\eta = (\nabla_\zeta \, \eta) \cdot g$ where the operator $\nabla_\zeta \, \eta$ is defined by

$$\nabla_\zeta \, \eta = -\frac{1}{2} \left\{ [\zeta,\eta]_{\mathscr{G}} - \mathbb{I}^{-1} (ad^*_\zeta \mathbb{I}\eta + ad^*_\eta \mathbb{I}\zeta) \right\}. \tag{7.192}$$

Thus

$$\omega^R(\zeta^R)\eta^R = -\frac{1}{2} \left\{ \mathrm{ad}_{\zeta^R} \eta^R - \mathbb{I}_R^{-1} \left(ad^*_{\zeta^R} \mathbb{I}_R \eta^R + ad^*_{\eta^R} \mathbb{I}_R \zeta^R \right) \right\}.$$

Since for $\dot{g} = \zeta^R \cdot g = g \cdot \zeta^L$

$$\omega^L(\dot{g}) = \mathrm{ad}_{\zeta^L} + \mathrm{Ad}_{g^{-1}} \omega^R(\zeta^R) \mathrm{Ad}_g.$$

we have that

$$\begin{aligned}
\omega^L(\zeta^L)\eta^L &= \mathrm{ad}_{\zeta^L}\eta^L + \mathrm{Ad}_{g^{-1}} \omega^R(\zeta^R) \mathrm{Ad}_g \eta^L \\
&= \mathrm{ad}_{\zeta^L}\eta^L - \frac{1}{2} \mathrm{Ad}_{g^{-1}} \left\{ \mathrm{ad}_{\zeta^R} \eta^R - \mathbb{I}_L^{-1} \left(ad^*_{\zeta^R} \mathbb{I}_R \eta^R + ad^*_{\eta^R} \mathbb{I}_R \zeta^R \right) \right\} \\
&= \mathrm{ad}_{\zeta^L}\eta^L - \frac{1}{2} \left\{ \mathrm{ad}_{\zeta^L} \eta^L - \mathbb{I}_L^{-1} \left(ad^*_{\zeta^L} \mathbb{I}_L \eta^L + ad^*_{\eta^L} \mathbb{I}_L \zeta^L \right) \right\} \\
&= \frac{1}{2} \left\{ \mathrm{ad}_{\zeta^L} \eta^L + \mathbb{I}_L^{-1} \left(ad^*_{\zeta^L} \mathbb{I}_L \eta^L + ad^*_{\eta^L} \mathbb{I}_L \zeta^L \right) \right\},
\end{aligned}$$

and hence that

$$\nabla_\xi \, \eta \triangleq d\eta(\xi) + \frac{1}{2} \left(\pm \mathrm{ad}_\xi \, \eta + \mathbb{I}^{-1} \left(\mathrm{ad}^*_\xi \mathbb{I}\eta + \mathrm{ad}^*_\eta \mathbb{I}\xi \right) \right),$$

where the $+$ results in $\nabla^L_{\xi^L} \eta^L$ and the $-$ results in $\nabla^R_{\xi^R} \eta^R$.

7.15.3 Summary: Levi-Civita Connection on a Lie Group

We conclude this section by pointing out that if the Riemannian metric satisfies certain invariant properties then the Levi-Civita connection can be explicitly expressed without the use of co-ordinates on G or $\mathscr{G}$. Specifically it can be shown that, for left-invariant metrics,

$$\nabla_{\xi}\eta \triangleq d\eta(\xi)+\frac{1}{2}\left(\pm \mathrm{ad}_{\xi}\eta - \mathbb{I}^{-1}\left(\mathrm{ad}^*_{\xi}\mathbb{I}\eta+\mathrm{ad}^*_{\eta}\mathbb{I}\xi\right)\right), \tag{7.193}$$

and for right-invariant metrics,

$$\nabla_{\xi}\eta \triangleq d\eta(\xi)+\frac{1}{2}\left(\pm \mathrm{ad}_{\xi}\eta + \mathbb{I}^{-1}\left(\mathrm{ad}^*_{\xi}\mathbb{I}\eta+\mathrm{ad}^*_{\eta}\mathbb{I}\xi\right)\right), \tag{7.194}$$

where the $+$ results in $\nabla^L_{\xi^L}\eta^L$ and the $-$ results in $\nabla^R_{\xi^R}\eta^R$. Here $d\eta(\xi)\triangleq \frac{d}{dt}\big|_{t=0}\eta(g\exp\xi t)$, for left-velocities while $d\eta(\xi)\triangleq \frac{d}{dt}\big|_{t=0}\eta((\exp\xi t)g)$, for right-velocities. A metric that is both left and right invariant is said to be a bi-invariant metric. It can be shown that for a bi-invariant metric $\left(\mathrm{ad}^*_{\xi}\mathbb{I}\eta+\mathrm{ad}^*_{\eta}\mathbb{I}\xi\right)\equiv 0$ and hence it follows that the bi-invariant connection is given by

$$\nabla_{\xi}\eta \triangleq d\eta(\xi)\pm\frac{1}{2}\mathrm{ad}_{\xi}\eta, \tag{7.195}$$

where once again the $+$ results in $\nabla^L_{\xi^L}\eta^L$ and the $-$ results in $\nabla^R_{\xi^R}\eta^R$.

7.15.4 The Symmetric Product

The symmetric product is defined by [**?**] to analyze controllability and accessibility properties of mechanical systems. The symmetric product between two vector fields $X=\xi^k(g)E_k$, $Y=\eta^k(g)E_k$ on G is defined to be,

$$<X:Y>=\nabla_X Y+\nabla_Y X. \tag{7.196}$$

Which from (7.190) results in,

$$\begin{aligned}<X:Y>=\nabla_X Y+\nabla_Y X=&\left\{d\eta^k(X)+d\xi^k(Y)\right\}E_k\\ &-gI^{-1}\left(ad^*_{\xi(g)}I\eta(g)+ad^*_{\eta(g)}I\xi(g)\right).\end{aligned} \tag{7.197}$$

Note that in the case of X and Y being left invariant vector fields on G, the first term in the right hand side of expression (7.197) vanishes and we have a purely algebraic operation on $\mathscr{G}$.

7.15.5 Curvature of a left invariant metric

Let u,v,w be tangent vectors to G at g and let the vector fields X,Y,Z be their arbitrary smooth extensions in a neighborhood of g. Then the curvature tensor on G is defined by [**?**, **?**, **?**, **?**],

$$R(u,v)w = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z. \tag{7.198}$$

The curvature also satisfies the following easily verifiable properties for any given vector fields X,Y,Z,W [**?**],

$$R(X,Y)Z = -R(Y,X)Z, \tag{7.199}$$

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0, \tag{7.200}$$

$$<< R(X,Y)Z\,,W >> = - << R(X,Y)W\,,Z >>, \tag{7.201}$$

$$<< R(X,Y)Z\,,W >> = << R(Z,W)X\,,Y >>, \tag{7.202}$$

In general (7.199) is true for any connection, (7.200) is true for any torsion free connection, (7.201) is true for any connection that is metric and (7.202) is true if the connection is both metric and torsion free. The condition (7.200) is referred to as the first Bianchi identity.

The sectional curvature of two vectors u,v is defined by,

$$K(u,v) = \frac{<< R(u,v)v\,,u >>}{<< u,u >><< v,v >> - << u,v >>^2}. \tag{7.203}$$

This quantity measures the infinitesimal rotation of any vector resulting from the parallel transport of the vector along an infinitesimally small parallelogram spanned by the vectors u,v [**?**]. Condition (7.202) gives that $K(u,v) = K(v,u)$. The entire curvature tensor is determined by the sectional curvature ([**?**], Lemma 3.3.3).

7.15.6 Cartan's Formulation

We reformulate the Levi-Civita connection and the curvature on G by means of the notion of vector valued forms. A general vector valued p-form is defined as $\alpha = E_k \otimes \alpha^k$, where all α^k are p-forms on G. Note that $\nabla_X E_j = \omega_{ij}^k X^i E_k = \omega_{ij}^k \sigma^i(X) E_k$. Thus ∇E_j can be considered as the vector valued 1-form given by,

$$\nabla E_j = E_k \otimes \omega_j^k, \tag{7.204}$$

where ω_j^k are the connection 1-forms given by

$$\omega_j^k = \omega_{ij}^k \sigma^i \tag{7.205}$$

and ω_{ij}^k is given by (7.188). Now it follows from (7.182) that,

$$\nabla Y = \nabla Y^j E_j = E_k \otimes (dY^k + \omega_{ij}^k \sigma^i Y^j). \tag{7.206}$$

7.15.6.1 Covariant Differentiation of Vector Valued Forms

A vector field Y can be thought of as a vector valued function. Then the vector valued 1-form ∇Y given by (7.206) can be thought of as the covariant derivative of Y. Generalizing this notion a general covariant differentiation operation that takes a vector valued p-form to a vector valued $p+1$-form is defined as follows (see [?] for details). For a vector valued p-form $\alpha = E_k \otimes \alpha^k$ the covariant derivative $\nabla\alpha$ is defined by requiring a Liebniz rule as follows.

$$\nabla\alpha = E_k \otimes d\alpha^k + \nabla E_k \otimes_\wedge \alpha_k\,, \tag{7.207}$$

where $\otimes_\wedge$ is defined as $\nabla E_k \otimes_\wedge \alpha^k = (E_r \otimes \omega^r_k) \otimes_\wedge \alpha^k := E_r \otimes \omega^r_k \wedge \alpha^k$. Thus

$$\nabla\alpha = E_k \otimes d\alpha^k + E_k \otimes \omega^k_j \wedge \alpha^j\,, \tag{7.208}$$

7.15.6.2 Curvature Two Forms

The vector valued curvature 2-form $\nabla\nabla E_j$ can now be defined as the covariant derivative of the connection 1-form ∇E_j. Explicitly we have

$$\nabla\nabla E_j = E_k \otimes (d\omega^k_j + \omega^k_r \wedge \omega^r_j)\,. \tag{7.209}$$

Therefore, the curvature 2-forms θ^k_j are given by

$$\theta^k_j = d\omega^k_j + \omega^k_r \wedge \omega^r_j. \tag{7.210}$$

Using (7.206) and (7.208), we also have that

$$\nabla\nabla Z = \nabla\nabla Z^j E_j = E_k \otimes \theta^k_j Z^j\,. \tag{7.211}$$

Note that $\nabla\nabla Z$ is linear in Z. Now let us compute the curvature 2-forms θ^k_j. Before we do this we recall the Maurer-Cartan structure equations

$$d\sigma^i = -\frac{1}{2} C^i_{ab} \sigma^a \wedge \sigma^b. \tag{7.212}$$

Recalling that the connection coefficients ω^k_{ij} given by (7.188) are constants we have that,

$$d\omega^k_j = w^k_{ij} d\sigma^i = -\frac{1}{2}\omega^k_{ij} C^i_{ab} \sigma^a \wedge \sigma^b, \tag{7.213}$$

and

$$\omega^k_r \wedge \omega^r_j = w^k_{ar}\sigma^a \wedge w^r_{bj}\sigma^b = \omega^k_{ar}\omega^r_{bj}\sigma^a \wedge \sigma^b. \tag{7.214}$$

Thus the curvature 2-forms are given by,

$$\theta^k_j = \frac{1}{2} R^k_{jab}\,\sigma^a \wedge \sigma^b \; = (-\frac{1}{2}\omega^k_{rj} C^r_{ab} + \omega^k_{ar}\omega^r_{bj})\,\sigma^a \wedge \sigma^b\,, \tag{7.215}$$

where the curvature coefficients R^k_{jab} are given by,

$$R^k_{jab} = (-\omega^k_{rj} C^r_{ab} + 2\omega^k_{ar}\omega^r_{bj}). \tag{7.216}$$

Therefore, at $g \in G$ if $\zeta, \eta, \xi \in \mathcal{G}$ are fixed, and X, Y, Z are arbitrary smooth vector fields defined in some neighborhood of g such that they take values $g \cdot \zeta, g \cdot \eta, g \cdot \xi$, respectively, at g then from (7.198) we have

$$\begin{aligned} R(X(g),Y(g))Z(g) &= 2E^k \otimes \theta^k_j(X,Y)Z^j - \nabla_{[X,Y]}Z, \\ &= E^k \otimes \left\{ 2\theta^k_j(X,Y)Z^j - dZ^k([X,Y]) - \omega^k_j([X,Y])Z^j \right\}. \end{aligned} \tag{7.217}$$

Note that since the extensions are arbitrary we could very well have picked them to be left invariant and then we have that,

$$R(g\cdot\zeta\,,\,g\cdot\eta)g\cdot\xi = E^k \otimes \{2\theta^k_j(g\cdot\zeta\,,\,g\cdot\eta)\xi^j - \omega^k_j(g\cdot[\zeta,\eta]_{\mathcal{G}})\xi^j\}. \tag{7.218}$$

Hence the pull back of the curvature to $\mathcal{G}$ by left translation is given by,

$$\begin{aligned} R(\zeta\,,\,\eta)\xi &= e^k \otimes \{2\theta^k_j(\zeta\,,\,\eta)\xi^j - \omega^k_j([\zeta,\eta]_{\mathcal{G}})\xi^j\}, \\ &= e^k\{R^k_{jab}\xi^j(\zeta^a\eta^b - \zeta^b\eta^a) - \omega^k_{ij}C^i_{ab}\zeta^a\eta^b\xi^j\}, \end{aligned} \tag{7.219}$$

where now θ^k_j and ω^k_j are understood to be the pull back of the curvature and connection forms to $\mathcal{G}$ under left translation.

7.15.7 Geodesics And Distances On A Lie Group

Let $\gamma(t)$ be a smooth curve on G. A vector field Y defined along $\gamma(t)$ is said to be parallely transported field along $\gamma(t)$ if $\nabla_{\dot{\gamma}}Y = 0$. Any smooth vector field Y along $\gamma(t)$ has the form $Y(t) = \gamma(t)\cdot\eta(t)$ for some smooth function $\eta(t) \in \mathcal{G}$. Thus, from (7.190), $Y(t) = \gamma(t)\cdot\eta(t)$ is parallely transported along $\gamma(t)$ if $\eta(t)$ satisfies,

$$\dot{\eta} = -\frac{1}{2}\left([g^{-1}\cdot\dot{\gamma},\,\eta] - I^{-1}\left(ad^*_{(g^{-1}\cdot\dot{\gamma})}I\eta + ad^*_{\eta}\,(I\,g^{-1}\cdot\dot{\gamma})\right)\right). \tag{7.220}$$

This is a linear time dependent ODE and it is known that for any given $\eta(0) = \eta_0$ there exists a unique solution, defined for all t for which $\gamma(t)$ is defined. Thus we know that there exists a unique parallely transported vector field $\gamma(t)\cdot\eta(t)$ along $\gamma(t)$ with initial condition η_0.

A smooth curve $\gamma(t)$ on G is a geodesic on G if $\dot{\gamma}(t)$ is parallely transported along $\gamma(t)$ or in other words if $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$. Noting that $\dot{\gamma}(t) = \gamma(t)\cdot\zeta(t)$, where $\zeta(t)$ is a smooth curve in $T_eG \simeq \mathcal{G}$, this results in the following ODE,

$$\dot{\zeta} = I^{-1}ad^*_{\zeta}I\zeta. \tag{7.221}$$

Thus geodesics on G are given by the solutions of the system of ODE's,

$$\dot{\gamma} = \gamma\cdot\zeta, \tag{7.222}$$
$$\dot{\zeta} = I^{-1}ad^*_{\zeta}\,I\zeta. \tag{7.223}$$

Local existence and uniqueness theorems for ODE's guarantee that given any $g_0 \in G$ and $\zeta_0 \in \mathscr{G}$ there exists a unique geodesic $\gamma(t) : (-\varepsilon,\varepsilon) \mapsto G$ such that $\gamma(0) = g_0$ and $\dot{\gamma}(t) = g_0 \cdot \zeta_0$ for some $\varepsilon > 0$. It also follows from these theorems that if $\gamma(t) : (-\varepsilon,\varepsilon) \mapsto G$ is a geodesic on G with $\gamma(0) = g$ and $\dot{\gamma}(0) = g \cdot \eta$ then there exists a unique geodesic through any $\tilde{g}$ sufficiently close to g with any initial velocity $\tilde{g} \cdot \tilde{\eta}$ sufficiently close to $g \cdot \eta$ and defined at least on $(-\varepsilon,\varepsilon)$. In addition on a Lie group we have the property that the geodesics are defined for all time. That is, a Lie group is geodesically complete. In what follows we provide a proof of this.

Note that for a geodesic curve $\gamma(t)$ the condition that the connection is metric yields,

$$\frac{d}{dt} << \dot{\gamma}, \dot{\gamma} >>= 2 << \nabla_{\dot{\gamma}}\dot{\gamma}, \dot{\gamma} >>= 0.$$

Hence $||\dot{\gamma}|| = \sqrt{< I\zeta, \zeta >} = \sqrt{I\zeta \cdot \zeta}$ is a constant along a geodesic curve $\gamma(t)$. If $||\dot{\gamma}|| = 1$ then the curve is said to be parameterized by arc length. Since $I\zeta \cdot \zeta$ is a constant along the trajectory of (7.223) and since $I\zeta \cdot \zeta$ is a positive definite quadratic form in ζ then a trajectory of (7.223) through $\zeta(0)$ at $t = 0$ lies entirely in the compact level set of $I\zeta \cdot \zeta$ through $\zeta(0)$ and hence we know that the solution is defined for all $t \in \mathscr{R}$. For a matrix Lie group this, in a straightforward sense, implies that the solutions of (7.222) – (7.223) are defined for all $t \in \mathscr{R}$. Thus matrix Lie groups are geodesically complete. Before we proceed to show this in general for any Lie group, we state and prove the following theorem.

Theorem 7.9. *If $\gamma(t)$ for $t \in (-\varepsilon,\varepsilon)$ is a geodesic through the identity element e of the Lie group G with $\dot{\gamma}(0) = \zeta_0$ for some $\varepsilon \in \mathscr{R}$, then $g\gamma(t)$ is the geodesic through g with initial velocity $g \cdot \zeta_0$.*

Let $\dot{\gamma}(t) = \gamma(t) \cdot \zeta(t)$ for some smooth curve $\zeta(t) \in \mathscr{G}$ with $\zeta(0) = \zeta_0$. By hypothesis $\gamma(t)$ is the geodesic through the identity e in G with initial velocity ζ_0. Hence $\gamma(t)$ is the solution of (7.222) – (7.223) with $\gamma(0) = e$ and $\zeta(0) = \zeta_0$. Consider $h(t) = g\gamma(t)$ where $h(0) = g$ and $\dot{h}(0) = g \cdot \zeta_0$. Note that $\dot{h}(t) = g\gamma(t) \cdot \zeta(t) = h(t) \cdot \zeta(t)$. Thus $h(t)$ is also a solution of (7.222) – (7.223) with $h(0) = g$ and $\zeta(0) = \zeta_0$ and hence is the unique geodesic through g with initial velocity $g \cdot \zeta_0$.

Using this theorem we can prove the following.

Theorem 7.10. *Any geodesic $\gamma(t)$ through the identity is defined for all $t \in \mathscr{R}$. Thus a Lie group G equipped with a left invariant metric is geodesically complete.*

From the existence and uniqueness of ODE's, given any $\zeta_0 \in \mathscr{G}$ there exists a unique geodesic $\gamma(t)$, defined on some maximal interval (α, β) with $\beta < \infty$, and passing through the identity with initial velocity ζ_0. Let $U = \{\zeta \in \mathscr{G} | I\zeta \cdot \zeta = I\zeta_0 \cdot \zeta_0\}$. The set U is compact. For any $\zeta \in U$ the existence and uniqueness theorems in ODE's imply that there exists some $\varepsilon_\zeta > 0$ and some U_ζ open in U such that there exists a unique geodesic through the identity of G defined at least on $[-\varepsilon_\zeta, \varepsilon_\zeta]$ with initial velocity $\zeta_0 \in U_\zeta$. The sets U_ζ for all $\zeta \in U$ form an open covering of U. Since U is compact there exists a finite subcover $U_{\tilde{\zeta}}$ of U. Let ε be the minimum of the $\varepsilon_{\tilde{\zeta}}$. Let $\tau = \beta - \varepsilon/2$. Consider the geodesic $\delta(t)$ passing through the identity with initial velocity $\gamma(\tau)^{-1} \cdot \dot{\gamma}(\tau)$. From Theorem 7.9 we have that the geodesic passing through $\gamma(\tau)$ with initial velocity $\dot{\gamma}(\tau)$ is $\gamma(\tau)\delta(t)$. However from the uniqueness of the solutions this should coincide with $\gamma(t)$ in the interval $[\tau, \beta)$ and hence by concatenating

$\gamma(t)$ with $\delta(t)$ in the interval $[\tau, \beta+\varepsilon/2)$ we have extended the domain of definition of $\gamma(t)$ by $\varepsilon/2$ thus contradicting the maximality of the domain of definition. Therefore the geodesic $\gamma(t)$ is defined for all $t \in \mathscr{R}$.

Note that if ζ is such that $ad^*_{\zeta} I\zeta = 0$ then $(\gamma(t), \zeta) = (\gamma(0)\exp(t\zeta), \zeta)$ is a solution of (7.222) – (7.223) with initial conditions $(\gamma(0), \zeta)$. Thus $\gamma(t) = \gamma(0)\cdot\exp(t\zeta)$ is a geodesic on G through $\gamma(0)$ with initial velocity $\gamma(0)\cdot\zeta$. At the identity they correspond to a 1-parameter subgroup of G and are a special class of geodesics among the very many. Also note that if the Lie group is abelian then the geodesics through the identity are exactly the 1-parameter sub groups of the Lie group.

7.15.8 The Local Distance Function on a Lie Group

On a Riemannian manifold, there exists the notion of a length of vectors at a point. Therefore, it is also possible to define the length of a piece wise smooth curve $\gamma(t)$ starting at $t = a$ at g_1 and ending at g_2 at $t = b$ as the integral of the magnitude of the velocity vector of the curve. If $l(\gamma)$ is the length of the curve, it will be explicitly given by,

$$l(\gamma) = \int_a^b << \dot{\gamma}, \dot{\gamma} >>^{\frac{1}{2}} dt. \tag{7.224}$$

Note that the value of this integral is independent of the parameterization of the curve. Given any two points g_1 and g_2 on a Riemannian manifold $(G, << \cdot, \cdot >>)$ define the set of curves,

$$\Lambda(g_1, g_2) = \{\gamma : [0\ 1] \mapsto G \mid \gamma \text{is piece wise smooth and}\, \gamma(0) = g_1, \gamma(1) = g_2\}. \tag{7.225}$$

Then the distance between g_1 and g_2 is defined as,

$$d(g_1, g_2) = \inf\{l(\gamma) : \gamma \in \Lambda(g_1, g_2)\}, \tag{7.226}$$

and defines a metric on the Riemannian manifold $(G, << \cdot, \cdot >>)$ [?, ?]. If a curve $\gamma \in \xi(g_1, g_2)$ exists such that $d(g_1, g_2) = l(\gamma)$ then it is referred to as a segment. It is known that segments are always geodesics. It is also known that the topology obtained from this metric is the same as the manifold topology (Theorem 3.4 of [?]). If the Riemannian manifold $(G, << \cdot, \cdot >>)$ equipped with the metric defined by (7.226) is complete as a metric space, then any two points on G can be joined by at least one segment. The Hopf-Rinow theorem (Theorem 7.1 of [?]) says that geodesic completeness and metric completeness are equivalent. Thus from Theorem 7.10 it also follows that any two points on a Lie group G can be joined by at least one segment. If the two points are sufficiently close then they are connected by a unique segment. A function $f : U \subset G \mapsto \mathscr{R}$, with U open, is a distance function if $||\nabla f|| \equiv 1$. Integral curves of ∇f are known to be segments (Lemma 3.6 of [?]). Furthermore any two points in U will be joined by an integral curve of $grad\, f$. For two sufficiently close points $\tilde{g}$ and g, there exists a unique $\zeta_e \in \mathscr{G}$ such that

$$\exp \zeta_e = \tilde{g} g^{-1}. \tag{7.227}$$

Let $U \subset G$ be an open set containing $g \in G$ such that there exists a unique ζ_e satisfying (7.227) for any $\tilde{g} \in U$. For a fixed $g \in G$ define the function

$$f(\tilde{g}) = ||\zeta_e||_{\mathscr{G}}. \tag{7.228}$$

Theorem 7.11. *The function $f(\tilde{g})$ defined by equation (7.228) is a local distance function on a Lie group.*

The gradient of $f(\tilde{g})$ evaluated at $\tilde{g}$ and denoted by $grad\, f(\tilde{g})$, is uniquely given by the relationship

$$<< grad\, f(\tilde{g})\ ,\ \tilde{g}\cdot\eta >>=< df\ ,\ \tilde{g}\cdot\eta >= \frac{I\zeta_e\cdot\eta}{\sqrt{I\zeta_e\cdot\zeta_e}} =< \frac{I\zeta_e}{\sqrt{I\zeta_e\cdot\zeta_e}}, \eta >, \tag{7.229}$$

for any $\eta \in \mathscr{G}$. Thus explicitly

$$grad\, f(\tilde{g}) = \tilde{g}\cdot(\frac{\zeta_e}{\sqrt{I\zeta_e\cdot\zeta_e}}). \tag{7.230}$$

Observe that $||grad\, f(\tilde{g})|| \equiv 1$. Hence f is a local distance function on G.

From Lemma 3.6 of [**?**], we thus have that

$$d(\tilde{g}, g) = f(\tilde{g}) = ||\zeta_e||_{\mathscr{G}}, \tag{7.231}$$

where $d(\tilde{g}, g)$ is the geodesic distance between g and $\tilde{g}$.

7.16 de Rham's Cohomology

Consider the set of points $P_0 = (0,\cdots,0), P_1 = (1,\cdots,0),\cdots,P_p = (0,\cdots,1)$ in $\mathbb{R}^p$. The convex set of points Δ_p contained in the polyhedron with vertices at $(P_0, P_1,\cdots,P_p)$ is called an *Euclidean p-simplex*. We will assign the orientation to Δ_p the orientation given by the 'vectors' $(P_1 - P_0), (P_2 - P_0), \cdots, (P_p - P_0)$.

A *Singular p-simplex* on a smooth manifold M^n is a map

$$\sigma : \Delta_p \to M^n.$$

The rank of this map need not be specified. The integration of a p-form will naturally be carried over σ. That is

$$\int_{\sigma(\Delta_p)} \alpha^p \triangleq \int_{\Delta_p} \sigma^* \alpha^p.$$

The k^{th} face of a Euclidean simplex Δ_p opposite the vertex P_k is denoted by the set of points $\Delta_{p-1}^{(k)} = (P_0, P_1,\cdots,P_{k-1},\widehat{P}_k,P_{k+1},\cdots,P_p)$. It is not a $(p-1)$-simplex since it is in $\mathbb{R}^p$ instead of $\mathbb{R}^{p-1}$. However it may be considered as a singular $(p-1)$-simplex in $\mathbb{R}^{p-1}$ by considering the map $f_k : \Delta_{p-1} \to \Delta_p$ defined by $P_0 \to P_0, \cdots, P_{k-1} \to P_{k-1}, P_k \to P_{k+1}, \cdots, P_{p-1} \to P_p$ and interior points of Δ_{p-1} to the interior of the k^{th} face.

The boundary of the Euclidean simplex Δ_p is defined by the formal sum

$$\partial \Delta_p = \sum_{k=0}^{p} (-1)^k \Delta_{p-1}^{(k)}. \tag{7.232}$$

In the case where $p = 2$ we see that $\partial \Delta_p$ is simply the counter clockwise oriented boundary of the triangle given by $\Delta_p = (P_0, P_1, P_2)$. Similarly one also sees that since $\partial \Delta_3 = (P_1, P_2, P_3) - (P_0, P_2, P_3) + (P_0, P_1, P_2) - (P_0, P_1, P_2)$ when $p = 3$ the sum $\partial \Delta_3$ represents the counter clockwise oriented sum of the faces. The object (7.232) is called an *integer* $(p-1)$ *chain*.

The idea of the integer p chain can be generalized to a *singular p chain* with coefficients in an abelian group G by defining

$$c_p = \sum_{s=1}^{m} g_s \sigma_p^s, \tag{7.233}$$

where $\sigma_p^s : \Delta_p \to M^n$ is a singular p-complex on M^n. The chain c_p can be considered to be a function on the space of singular complexes of M^n with values in G by defining $c_p(\sigma_p^s) = g^s$. In particularly if σ_p is not in the formal sum of (7.233) then $c_p(\sigma_p) = 0$.

Defining $c_p(\sigma^s) + c_p{}'(\sigma^s) = g^s + g^{s\prime}$ we can give a group structure to the space of all singular p-chains, $C_p(M^n; G)$. Then the Euclidean p-simplex can be considered as an element of $C_p(\mathbb{R}^p; \mathbb{Z})$. Define the operator

$$\partial : C_p(M^n; G) \to C_{p-1}(M^n; G),$$

by the relationship $\partial \sigma \triangleq \sigma^*(\partial \Delta_p)$ which gives

$$\partial c_p = \sum_{s=1}^{m} g^s \partial \sigma_p^s.$$

Thus ∂ is a group homomorphism. Furthermore one can show that $\partial^2 = 0$. A p-chain such that its boundary is zero will be called a p-cycle.

Define

$$Z_p(M^n; G) = \ker\ \partial : C_p \to C_{p-1}, \tag{7.234}$$

$$B_p(M^n; G) = \operatorname{Im}\ \partial : C_{p+1} \to C_p. \tag{7.235}$$

Both are sub groups of C_p. Define

$$H_p(M^n; G) = \frac{Z_p(M^n; G)}{B_p(M^n; G)}. \tag{7.236}$$

That is the equivalent class of p-cycles that differ by a p-chain that is the boundary of a $p+1$ chain. Specifically if $z_p, z_p' \in Z_p(M^n; G)$ then $z_p - z_p' = \partial c_{p+1}$ for some $p+1$ chain.

The space $H_p(M^n;G)$ is called the p^{th} *Homology Group*. The p^{th} *Betti number* of M^n is defined to be $b_p(M) = \dim(H_p(M^n;\mathbb{R}))$. That is it is the maximal number of p-cycles that the linear combination of any of them is never a boundary of a $(p+1)$-chain.

We see that if α is a closed p-form then

$$\int_{z_p} \alpha - \int_{z'_p} \alpha = \int_{z_p - z'_p} \alpha = \int_{\partial c_{p+1}} \alpha = \int_{c_{p+1}} d\alpha = 0.$$

Thus the integral of a closed p-form over any two equivalent p-cycles is the same. On the other hand if α and β are two closed p-forms such that $\alpha - \beta = d\gamma$ and hence is exact then

$$\int_{z_p} \alpha - \int_{z_p} \beta = \int_{z_p} (\alpha - \beta) = \int_{z_p} d\gamma = \int_{\partial z_p} \gamma = \int_0 \gamma = 0,$$

and thus the value of the integral of two closed forms that differ by an exact form, over any cycle, is also the same. Thus one can uniquely identify closed p-forms that differ by an exact form with p-cycles that differ by the boundary of a $(p+1)$-chain. Let us state this formally by defining

$$F^p = \ker\ d : \wedge^p \to \wedge^{p+1}, \tag{7.237}$$

$$E^p = \mathrm{Im}\ d : \wedge^{p-1} \to \wedge^p. \tag{7.238}$$

and the space of equivalence classes

$$\mathscr{R} = \frac{F^p}{E^p}.$$

The map given by the relationship

$$\mathscr{I}([\alpha])([z]) \triangleq \int_z \alpha,$$

is map $\mathscr{I} : \mathscr{R} \to H_p^*(M^n;\mathbb{R})$ since from the previous discussion it follows that this value is independent of the representation of α for $[\alpha]$ and z for $[z]$. The dual space $H_p^*(M^n;\mathbb{R})$ is called the p^{th} *Cohomology Group* of M^n. In 1931 de Rham showed that this map is an isomorphism.

Theorem 7.12. : de Rham's Theorem: *Consider the map $\mathscr{I} : \mathscr{R} \to H_p^*(M^n;\mathbb{R})$.*

(i) If b_p is the p^{th} Betti number of M^n, $\{z_p^{(1)}, \cdots, z_p^{(b_p)}\}$ is basis of cycles for the Homology group $H_p(M^n;\mathbb{R})$, and $\pi_1, \cdots, \pi_{b_p}$ are real constants, then there exists a closed p-form α^p such that

$$\int_{z_p^{(i)}} \alpha^p = \pi_i.$$

This shows that the map $\mathscr{I}$ is onto.

(ii) The map $\mathscr{I}$ is 1:1. Thus the kernel of $\mathscr{I}$ is $[0]$ and therefore if α^p is closed and $\int_{z_p^{(i)}} \alpha^p = 0$ for all cycles $z_p^{(i)}$ then α^p is equivalent to $[0] \in \mathscr{R}$ and thus α^p is exact. That is

$$\alpha^p = d\beta^{p-1}.$$