MAGNETIC SPHERICAL PENDULUM

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

SELMA YILDIRIM

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

THE DEPARTMENT OF MATHEMATICS

JULY 2003

Approval of the Graduate School of Natural and Applied Sciences.

Prof. Dr. Tayfur Öztürk
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Ersan Akyıldız
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Cem Tezer
Supervisor

Examining Committee Members

Prof. Dr. Hurşit Önsiper

Prof. Dr. Cem Tezer

Prof. Dr. Sinan Sertöz

Assist. Prof. Dr. Muhittin Uğuz

Assist. Prof. Dr. Ebru Keyman

ABSTRACT

MAGNETIC SPHERICAL PENDULUM

Yildirim, Selma

M. Sc., Department of Mathematics

Supervisor: Prof. Dr. Cem Tezer

July 2003, 49 pages

The magnetic spherical pendulum is a mechanical system consisting of a pendulum whereof the bob is electrically charged, moving under the influence of gravitation and the magnetic field induced by a magnetic monopole deposited at the origin. Physically not directly realizable, it turns out to be equivalent to a reduction of the Lagrange top. This work is essentially the logbook of our attempts at understanding the simplest contemporary approaches to the magnetic spherical pendulum.

Keywords: Magnetic spherical pendulum, Hamiltonian systems, Poisson brackets.

ÖZ

MANYETIK KÜRESEL SARKAÇ

Yildirim, Selma

Yüksek Lisans, Matematik Bölümü

Tez Yöneticisi: Prof. Dr. Cem Tezer

Temmuz 2003, 49 sayfa

Manyetik küresel sarkaç, ucundaki kütlesi aynı zamanda elektrikle yüklü olup, yerçekimi ve merkeze yerleştirilmiş tek bir manyetik kutup tarafindan üretilen manyetik alanın tesiri altında hareket eden sarkaçtan ibaret bir mekanik sistemdir. Gerçek bir fizik olayı olarak ortaya çıkmasa da, aslında Lagrange topaçının indirgenmiş şekline eşdeğer olduğu görülür. Bu çalısma esas itibarıyla, manyetik küresel sarkaça en basit çağdas yaklasımları anlama gayretlerimizin kayıtlarından meydana gelmektedir.

Anahtar Kelimeler: Manyetik küresel sarkaç, Hamilton sistemleri, Poisson parantezleri.

To lovely Zümrüt

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Prof. Dr. Cem Tezer for his patient guidance in the preperation of this thesis.

My thanks go to Türkay Yolcu as well, for the motivating and encouraging correspondance and assistance at various stages of my thesis.

I am very thankful to my family for their love, constant support and encouragement throughout my education especially to dear sister Zümrüt for her consideration and valuable advice.

I would like to thank all my friends for their kindness and camaraderie, especially to Arzu Baysal, Öznur Mut and Adem Çakmak.

TABLE OF CONTENTS

CHAPTER 1

SPHERICAL PENDULUM AND ITS MAGNETIC VERSION IN NAIVE TREATMENT

In classical mechanics, the term spherical pendulum refers to the configuration consisting of a massive particle suspended by a rigid weightless rod from a fixed point and subject to constant uniform gravitational field. Equivalently the spherical pendulum may be regarded as a particle constrained to move on a sphere under constant uniform gravitational field. This is an old and established problem of classical mechanics which has been subjected to very detailed treatment in numerous textbooks ([Whi],[Pars]). It is known that a full quantitive solution of the governing equations requires the use of elliptic functions.

An interesting if rather artificial modification of the problem ensues if the massive particle is also understood to be electrically charged and a magnetic monopole is placed at the center of the sphere on which the particle in question is constrained to move. In this case the particle is subject not only to the gravitation and the forces of constraint but also to the Lorentz force that acts on charged particles that move in magnetic fields. Clearly rather far fetched from a scientific point of view, the problem has been concocted for the purpose of illustrating the use of certain recent mathematical artifacts.

Classical treatments which sometimes involve multivalued coordinate systems and occasionally force the investigator to give separate treatments of non-generic orbits are in many respects unsatisfactory from a modern point of view, yet they are very pleasant mathematically and intuitively satisfying. We find it important first to go through such a hands-on investigation of the above mentioned problems to gain insight into the portent of a mathematical mechanism that clearly originates from similar physical problems. In fact we adopt an even more extremist stand in that we approach the phenomenon by direct analysis of the forces acting upon the particle, hence the adjective "naive" in our title. The position vector is

$$\mathsf{x} = \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} a\sin\theta\cos\varphi \\ a\sin\theta\sin\varphi \\ -a\cos\theta \end{bmatrix}$$

The linear momentum of the particle is $\mathsf{p} = m\dot{\mathsf{x}}$ and

$$\mathsf{p} = m\dot{\mathsf{x}} = ma(\begin{bmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ \sin\theta \end{bmatrix} \dot{\theta} + \begin{bmatrix} -\sin\theta\sin\varphi \\ \sin\theta\cos\varphi \\ 0 \end{bmatrix} \dot{\varphi})$$

The only forces acting on the particle are the gravitational force and the constraining force that arises from the tension τ in the rod. Consequently

$$\mathsf{F} = \begin{bmatrix} -\tau\sin\theta\cos\varphi \\ -\tau\sin\theta\sin\varphi \\ \tau\cos\theta - mg \end{bmatrix}$$

According to the second law of Newton, the motion of the particle can be described by the differential equation

$$\mathsf{F} = \frac{d\mathsf{p}}{dt}$$

Explicitly,

$$\begin{bmatrix} -\tau\sin\theta\cos\varphi \\ -\tau\sin\theta\sin\varphi \\ \tau\cos\theta - mg \end{bmatrix} = ma(\begin{bmatrix} -\sin\theta\cos\varphi \\ -\sin\theta\sin\varphi \\ \cos\theta \end{bmatrix} \dot{\theta}^2 + 2 \begin{bmatrix} -\cos\theta\sin\varphi \\ \cos\theta\cos\varphi \\ 0 \end{bmatrix} \dot{\theta}\,\dot{\varphi}$$

$$
+ \begin{bmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ \sin\theta \end{bmatrix} \ddot{\theta} + \begin{bmatrix} -\sin\theta\cos\varphi \\ -\sin\theta\sin\varphi \\ 0 \end{bmatrix} \dot{\varphi}^2
$$

$$
+ \begin{bmatrix} -\sin\theta\sin\varphi \\ \sin\theta\cos\varphi \\ 0 \end{bmatrix} \ddot{\varphi}) \tag{1.1}
$$

Upon taking inner product of both sides with $\begin{bmatrix} \sin\varphi \\ -\cos\varphi \\ 0 \end{bmatrix}$ we get

$$
0 = ma(-2\cos\theta\ \dot{\theta}\ \dot{\varphi} - \ \sin\theta\ddot{\varphi}) \tag{1.2}
$$

or

$$
2\dot{\theta}\dot{\varphi}\cos\theta + \sin\theta\ddot{\varphi} = 0
$$

and multiplying with $\sin\theta$

$$
\frac{d}{dt}(\dot{\varphi}\sin^2\theta) = 0
$$

equivalently

$$
\dot{\varphi}\sin^2\theta = \frac{L_z}{ma^2}
$$

where L_z is a constant that may be identified with z-component of the angular momentum of the particle with respect to the origin.

Upon taking inner product with

$$
\begin{bmatrix} \cos\varphi\cos\theta \\ \sin\varphi\cos\theta \\ \sin\theta \end{bmatrix}
$$

we obtain

$$
-mg\sin\theta = ma(\ddot{\theta} - \sin\theta\cos\theta\dot{\varphi}^2) \tag{1.3}
$$

Dividing out ma and multiplying with $\dot{\theta}$

$$
\dot{\theta}\ddot{\theta} - \sin\theta\cos\theta\dot{\theta}\dot{\varphi}^2 = -\frac{g}{a}\sin\theta\dot{\theta}
$$

and putting

$$\dot{\theta}\dot{\varphi}\cos\theta = -\frac{1}{2}\sin\theta\ddot{\varphi}$$

$$\dot{\theta}\ddot{\theta} + \frac{1}{2}\sin^2\theta\dot{\varphi}\ddot{\varphi} = -\frac{g}{a}\sin\theta\dot{\theta}$$

replacing $\sin^2\theta\dot{\varphi}$ with

$$\frac{L_z}{ma^2}$$

$$\dot{\theta}\ddot{\theta} + \frac{L_z}{2ma^2}\ddot{\varphi} = -\frac{g}{a}\sin\theta\dot{\theta}$$

equivalently

$$\frac{d}{dt}\left(\frac{1}{2}\dot{\theta}^2 + \frac{L_z}{2ma^2}\dot{\varphi}\right) = \frac{d}{dt}\left(\frac{g}{a}\cos\theta\right)$$

and again using $\sin^2\theta\dot{\varphi} = \dfrac{L_z}{ma^2}$ we find

$$\frac{d}{dt}(\frac{1}{2}\dot{\theta}^2 + \frac{1}{2}\dot{\varphi}^2\ \sin^2\theta) = \frac{d}{dt}\frac{g}{a}\cos\theta$$

Thus,

$$\frac{1}{2}(\dot{\theta}^2 + \dot{\varphi}^2\sin^2\theta) - \frac{g}{a}\cos\theta = constant$$

As $|\mathsf{p}|^2 = m^2a^2(\dot{\theta}^2 + \sin^2\dot{\varphi}^2)$ the left hand side can be seen to be E/ma^2 where E is the total energy of the system given by

$$E = \frac{|\mathsf{p}|^2}{2m} - mga\cos\theta$$

Before proceeding further, notice the particular case in which $L_z = 0$, which implies that φ is a constant and therefore the motion takes place in a fixed vertical plane. Hence,

$$\frac{1}{2}ma^2\dot{\theta}^2 - mga\cos\theta = E$$

upon differentiation

$$ma^2\ddot{\theta} + mga\sin\theta = 0$$

or equivalently ,

$$\ddot{\theta} = -\frac{g}{a}\sin\theta$$

which is the equation of motion for the planar pendulum.

In general the equations of motion read

$$\begin{aligned} \frac{1}{2}ma^2(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2\theta) - mga\cos\theta &= E \\ ma^2\dot{\varphi}\sin^2\theta &= L_z \end{aligned}$$

where E, L_z are constants corresponding to the total energy and the z-component of the angular momentum about the origin.

It is interesting to stop here and see how the equations (1.2) and (1.3) are modified in the presence of a magnetic monopole of strength μ at the origin and if the electric charge q is attached to the material particle. The magnetic induction resulting from the monopole is of the form

$$\beta = \mu\frac{\mathsf{x}}{\|\mathsf{x}\|^3} = \frac{\mu}{a^3}\mathsf{x}$$

This causes the Lorentz force

$$\mathsf{F} = q\dot{\mathsf{x}} \times \beta = \frac{q\mu}{a^3}\,\dot{\mathsf{x}} \times \mathsf{x}$$

to act on the particle. Explicitly, the Lorentz force is

$$\begin{aligned} &\frac{q\mu}{a^3}\left(\begin{bmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ \sin\theta \end{bmatrix}\dot{\theta} + \begin{bmatrix} -\sin\theta\sin\varphi \\ \sin\theta\cos\varphi \\ 0 \end{bmatrix}\dot{\varphi}\right) \times \begin{bmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ -\cos\theta \end{bmatrix} \\ &= \frac{q\mu}{a^3}\left(\begin{bmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{bmatrix}\dot{\theta} - \begin{bmatrix} \cos\theta\sin\theta\cos\varphi \\ \cos\theta\sin\theta\sin\varphi \\ \sin^2\theta \end{bmatrix}\dot{\varphi}\right) \end{aligned}$$

Adding the Lorentz force to the left hand side of (1.1) and taking inner product of this modified (1.1) with

$$\begin{bmatrix} \sin\varphi \\ -\cos\varphi \\ 0 \end{bmatrix}$$

we obtain the modified form of (1.2) as

$$\frac{q\mu}{a^3}\dot{\theta} = ma(-2\cos\theta\dot{\theta}\dot{\varphi} - \sin\theta\ddot{\varphi}) \tag{1.4}$$

and taking inner product with

$$\begin{bmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ \sin\theta \end{bmatrix}$$

we obtain the modified form of (1.3) as

$$-mg\sin\theta - \frac{q\mu}{a^3}\sin\theta\dot{\varphi} = ma(\ddot{\theta} - \sin\theta\cos\theta\dot{\varphi}^2) \tag{1.5}$$

We observe that (1.4) can be multiplied with $\dot{\theta}$ and written also in the form

$$\frac{d}{dt}\{ma^2\sin^2\theta\ \dot{\varphi} + \frac{q\mu}{a^3}\cos\theta\} = 0. \tag{1.6}$$

Clearly the quantity within the brackets is a conserved quantity . However, it may no longer be identified with the z-component of the angular momentum about the origin.

CHAPTER 2

NAIVE THEORY OF THE LAGRANGE TOP

Given a material particle with position vector x and mass m, moving under the action of a force F, the momentum $\mathsf{p} = m\dot{\mathsf{x}}$ of the particle is related to F by *Newton's Second Law*:

$$\frac{d}{dt}\mathsf{p} = \mathsf{F}.$$

Similarly the torque $\mathsf{N} = \mathsf{x} \times \mathsf{F}$ of the force in question is related to the angular momentum $\mathsf{L} = \mathsf{x} \times \mathsf{p}$ of the particle by

$$\begin{aligned} \frac{d}{dt}\mathsf{L} &= \frac{d\mathsf{x}}{dt} \times \mathsf{p} + \mathsf{x} \times \frac{d\mathsf{p}}{dt} \\ &= 0 + \mathsf{x} \times \mathsf{F} = \mathsf{N} \end{aligned}$$

Given a system consisting of finitely many material points with position vectors x_k and masses m_k where $k = 1, 2, 3, \ldots$, acted upon by forces F_k the total torque is $\mathsf{N} = \sum \mathsf{N}_k$ and the total angular momentum is $\mathsf{L} = \sum \mathsf{L}_k$ where L_k is the angular momentum of the kth particle and $\mathsf{N}_k = \mathsf{x}_k \times \mathsf{F}_k$ is the torque due to force F_k acting on the kth particle. These quantities are obviously related by

$$\frac{d\mathsf{L}}{dt} = \mathsf{N}.$$

We shall understand a rigid body as a limiting form of a system of finitely many points so that

$$L = \int_{Body} \mathsf{x} \times \rho(\mathsf{x})\dot{\mathsf{x}}\; dVol(\mathsf{x})$$

where ρ represents the mass density of the rigid body. Note that up to this point we have only effected a passage from the "discrete" to the "continuous" and made no use of rigidity assumption. From this point on we shall make the assumption that the system of points constitutes a rigid body with one point thereof is fixed at the origin. This assumption can be made to bear upon our considerations by writing $\mathsf{x} = \mathsf{x}(t)$ in the form

$$\mathsf{x} = \Omega \mathsf{x}(0)$$

where $\Omega = \Omega(t) \in SO(3)$. With this observation we obtain

$$\frac{d\mathsf{x}}{dt} = \Omega \frac{d\Omega}{dt} \mathsf{x}_0$$

On the other hand as $\Omega\Omega^T = 1$ we have

$$\begin{aligned} 0 = \frac{d}{dt}(\Omega\Omega^T) &= \frac{d\Omega}{dt}\Omega^T + \Omega^T \frac{d\Omega^T}{dt} \\ &= \frac{d\Omega}{dt}\Omega^T + (\frac{d\Omega}{dt}\Omega^T)^T \end{aligned}$$

equivalently

$$\Xi = \frac{d\Omega}{dt}\Omega^T \in so(3)$$

that is Ξ is a skew symmetric matrix.

We observe that each $A \in so(3)$ is of the form

$$A = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}$$

and for each $\mathsf{r} \in \mathbb{R}^3$

$$A\mathsf{r} = \begin{bmatrix} -c \\ b \\ -a \end{bmatrix} \times \mathsf{r}.$$

Let $i : so(3) \to \mathbb{R}^3$ be defined by

$$i(\begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}) = \begin{bmatrix} -c \\ b \\ -a \end{bmatrix}$$

We have already noted that

$$A\mathsf{r} = i(A) \times \mathsf{r}$$

for any $\mathsf{r} \in \mathbb{R}^3$.

Going back to the equation for the angular momentum we put

$$\omega = \omega(t) = i(\frac{d\Omega}{dt}\Omega^T)$$

The vector ω is called the *angular velocity* of the rigid body.

We obtain

$$\begin{aligned} \frac{d}{dt}\mathsf{x} &= \frac{d\Omega}{dt}\mathsf{x}_0 = \frac{d\Omega}{dt}\Omega^{-1}\mathsf{x} \\ &= \omega \times \mathsf{x} \end{aligned}$$

and

$$\begin{aligned} L &= \int_{Body} \mathsf{x} \times \rho(\mathsf{x})[\omega \times \mathsf{x}]dVol(\mathsf{x}) \\ &= \int_{Body} \rho(\mathsf{x})[x^T x\omega - x^T\omega x]dVol(\mathsf{x}) \\ &= I\omega \end{aligned}$$

where

$$I = \int_{Body} [x^T x 1 - xx^T]\rho(\mathsf{x})dVol(\mathsf{x})$$

which is a symmetric matrix referred to as the *inertia tensor* of the body. Presented in this form the matrix I is not useful since it is time dependent. This confronts us with a peculiarity of the rigid body dynamics that the equations of motion tend to be unmanagable unless they are written in reference to a system of reference that is fixed in the rigid body.

Quite generally a 3×1 matrix

$$\mathsf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = x\mathsf{e}_1 + y\mathsf{e}_2 + z\mathsf{e}_3$$

where

$$\mathsf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathsf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \mathsf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

is understood to consists of components $x_1,\ x_2,\ x_3$ of a vector in $\mathbb{R}^3$ with respect to a coordinate system xyz. This system is understood to be an inertial system, that is, with respect to it the laws of Newton are valid.

Let the components of the 3×1 matrix

$$\mathsf{y}_b = \begin{bmatrix} y_{1b} \\ y_{2b} \\ y_{3b} \end{bmatrix}$$

consist of components of the same vector with respect to a coordinate system $x_b y_b z_b$ fixed in the rigid body. Let $\mathsf{e}_{1b},\ \mathsf{e}_{2b},\ \mathsf{e}_{3b}$ be the unit vectors along the x_b-, y_b-,z_b-axes. Obviously there exists $\Omega \in SO(3)$ such that

$$\mathsf{e}_{kb} = \Omega \mathsf{e}_k$$

and

$$y_{kb} = \mathsf{e}_{kb} y = \mathsf{e}_{kb}^T y = (\Omega \mathsf{e}_k)^T y = \mathsf{e}_k \Omega^T y = (\Omega^T y) \mathsf{e}_k$$

for $k = 1,\ 2,\ 3$. We conclude that

$$\mathsf{y}_b = \Omega^{-1} \mathsf{y}.$$

Let us assume without loss of generality that the coordinate system $x_b y_b z_b$ fixed in the body coincides with the inertial frame xyz at $t = 0$. Let us now compute the inertia tensor of the rigid body in question with respect to the coordinate system fixed in the rigid body:

$$\begin{aligned} I_b &= \int_{Body} [x_b^T x_b 1 - x_b x_b^T] \rho(\mathsf{x}_b) dVol(\mathsf{x}_b) \\ &= \int_{\Omega^{-1}(Body)} [(\Omega^T x)^T (\Omega^T x) \not\Vdash - (\Omega^T x)(\Omega^T x)^T] \rho((\Omega^T x)) dVol(\Omega^T x) \end{aligned}$$

$$\mathsf{N}_b = \Omega^{-1}\mathsf{N} = \Omega^{-1}\frac{dL}{dt} = \Omega^{-1}I\omega = \Omega^{-1}I\Omega\omega_b = I_b\omega_b$$

It is an interesting exercise in advanced calculus to show that I_b is a constant matrix.

We now have

$$\begin{aligned}
\mathsf{N}_b &= \Omega^{-1}\mathsf{N} \\
&= \Omega^{-1}\frac{d\mathsf{L}}{dt} \\
&= \Omega^{-1}\frac{d}{dt}(I\omega) \\
&= \Omega^{-1}\frac{d}{dt}(\Omega I_b\omega_b) \\
&= \Omega^{-1}[\frac{d\Omega}{dt}I_b\omega_b + \Omega I_b\frac{d\omega_b}{dt}] \\
&= \Omega^{-1}[\omega \times (\Omega I_b\omega_b) + \Omega I_b\frac{d\omega_b}{dt}] \\
&= (\Omega^{-1}\omega) \times I_b\omega_b + I_b\dot{\omega}_b \\
&= \omega_b \times I_b\omega_b + I_b\dot{\omega}_b
\end{aligned}$$

Making, from this point on, the convention that all quantities are with reference to a coordinate system fixed in the body and dropping the subscript "b" we obtain the Euler equations

$$I\dot{\omega} + \omega \times I\omega = \mathsf{N}$$

Now that we have dropped subscripts, we have

$$\omega = \Omega^{-1}i(\frac{d}{dt}\Omega\Omega^{-1})$$

Further assuming without loss of generality that with respect to this coordinate system the inertia matrix has the diagonal form

$$I = \begin{bmatrix} I_1 & & \\ & I_2 & \\ & & I_3 \end{bmatrix}$$

we obtain the following explicit system of ordinary differential equations.

$$\begin{aligned}
I_1\dot{\omega}_1 - (I_2 - I_3)\omega_2\omega_3 &= N_1 \\
I_2\dot{\omega}_2 - (I_3 - I_1)\omega_3\omega_2 &= N_2 \\
I_3\dot{\omega}_3 - (I_1 - I_2)\omega_1\omega_2 &= N_3
\end{aligned}$$

Beatiful as the Euler equations are their intuitive content becomes clearer only after the introduction of the Euler angles. Let $Rot(k, \alpha) \in SO(3)$ denote the rotation about the directed line k through the angle $\alpha \in \mathbb{R}/2\pi\mathbb{Z}$. Each $\Omega \in SO(3)$ sufficiently close to the identity can be written in the form

$$\Omega = Rot(z'', \psi) \circ Rot(x', \theta) \circ Rot(z, \varphi)$$

where z denotes the z-axis, x' denotes the image of the x-axis under $Rot(z, \varphi)$ and z'' denotes the image of the z-axis under $Rot(x', \theta) \circ Rot(z, \varphi)$.
Clearly

$$Rot(z, \varphi) = R_3(\varphi)$$

where

$$\begin{aligned}
R_3(\varphi) &= \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
Rot(x', \theta) &= R_3(\varphi)R_1(\theta)R_3(\varphi)
\end{aligned}$$

where

$$R_1(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

and finally

$$Rot(z'', \psi) = [R_3(\varphi) \circ R_1(\theta)]R_3(\psi)[R_3(\varphi) \circ R_1(\theta)]^{-1}$$

By a simple regrouping, we obtain

$$\Omega = R_3(\varphi)R_1(\theta)R_3(\psi)$$

$$= \begin{bmatrix} \cos\varphi\cos\psi - \sin\varphi\sin\psi\cos\theta & -\cos\varphi\sin\psi - \sin\varphi\cos\psi\cos\theta & \sin\varphi\sin\theta \\ \sin\varphi\cos\psi + \cos\varphi\sin\psi\cos\theta & -\sin\varphi\sin\psi + \cos\varphi\cos\psi\cos\theta & -\cos\varphi\sin\theta \\ \sin\psi\sin\theta & \cos\psi\sin\theta & \cos\theta \end{bmatrix}$$

Let us reconsider the map $i : so(3) \to \mathbb{R}^3$ introduced above and note that for any $A \in SO(3)$ and $\Xi \in so(3)$ we have $A\Xi A^{-1} \in so(3)$ and

$$i(A\Xi A^{-1}) = Ai(\Xi).$$

There seems to be no elegant proof of this fact. The best method is to consider the map

$$\varphi : so(3) \longrightarrow so(3)$$

defined by

$$\varphi(A)(X) = A^{-1}i(Ai^{-1}(X)A^{-1})$$

and observe that

$$T_e\varphi : T_e so(3) \approxeq so(3) \longrightarrow T_e so(3)$$

is the identity map hence φ has to be the identity, too.

$$\begin{aligned}
\omega &= \Omega^{-1}i(\frac{d\Omega}{dt}\Omega^{-1}) \\
&= i(\Omega^{-1}\frac{d\Omega}{dt}) \\
&= i(R_3(-\psi)R_1(-\theta)R_3(-\varphi)[R_3'(\varphi)R_1(\theta)R_3(\psi)\dot{\varphi} + R_3(\varphi)R_1'(\theta)R_3(\psi)\dot{\theta} \\
&+ R_3(\varphi)R_1(\theta)R_3'(\psi)\dot{\psi}]) \\
&= \dot{\varphi}R_3(-\psi)R_1(-\theta)i(R_3(-\varphi)R_3'(\varphi)) + \dot{\theta}R_3(-\psi)i(R_1(-\theta)R_1'(\theta)) \\
&+ \dot{\psi}i(R_3(-\psi)R_3'(\psi)) \\
&= \dot{\varphi}R_3(-\psi)R_1(-\theta)\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + \dot{\theta}R_3(-\psi)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \dot{\psi}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\
&= \dot{\varphi}\begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
\end{aligned}$$

$$+ \quad \dot{\theta} \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \dot{\psi} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Consequently,

$$\begin{aligned} \omega_1 &= \dot{\varphi} \sin\theta \sin\psi + \dot{\theta} \cos\psi \\ \omega_2 &= \dot{\varphi} \sin\theta \cos\psi + \dot{\theta} \sin\psi \\ \omega_3 &= \dot{\varphi} \cos\theta + \dot{\psi} \end{aligned}$$

To complete our derivation of the governing equations we still have to compute the total torque of the forces action upon the particles in the rigid body. Once again reverting to our approximate picture of the rigid body as a system of finitely many material particles we assume that the ith material particle is acted upon by an "external" force $\mathsf{F}_i = m_i g \mathsf{e}_3$ due to gravitation and "internal" force f_{ij} due to the jth particle $j \neq i$ which obey the third Law of Newton. Thus, we assume that

$$\mathsf{f}_{ij} = a_{ij}(\mathsf{x}_j - \mathsf{x}_i)$$

for $i \neq j$ where $a_{ij} \in \mathbb{R}$ and $a_{ij} = a_{ji}$.
Consequently

$$\begin{aligned} \mathsf{N} &= \sum_i \mathsf{x}_i \times [m_i g \mathsf{e}_3 + a_{ij}(\mathsf{x}_j - \mathsf{x}_\mathsf{i})] \\ &= g(\sum_i m_i \mathsf{x}^i) \times \mathsf{e}_3 + \sum_i a_{ij} \mathsf{x}_i \times \mathsf{x}_j \\ &= g(\sum_i m_i \mathsf{x}_i) \times \mathsf{e}_3 \end{aligned}$$

Passing to the limit

$$\begin{aligned} \mathsf{N} &= g(\int_{Body} \mathsf{x}\rho(\mathsf{x}) dVol(\mathsf{x})) \times \mathsf{e}_3 \\ &= Mg\bar{\mathsf{x}} \times \mathsf{e}_3 \end{aligned}$$

where M is the total mass, $\overline{x}$ is the position vector of the center of mass of the rigid body.

Finally we restrict our attention to the special case called the "Lagrange Top", by making the following assumptions:

1. The body has rotational symmetry about the z_b-axis, i.e. the third principal axis, as a consequence of which $I_1 = I_2$.

2. The center of mass of the body lies on the third principal axis. That is

$$\bar{\mathsf{x}} = \mu\Omega\mathsf{e}_3.$$

Therefore

$$\mathsf{N} = Mg\mu(\Omega\mathsf{e}_3) \times \mathsf{e}_3$$

and

$$\mathsf{N}_b = Mg\mu\mathsf{e}_3 \times \Omega^{-1}\mathsf{e}_3.$$

Again we drop the subscript and refer everything to the principal axes fixed in the body and find

$$\begin{aligned}
\mathsf{N} &= Mg\mu\mathsf{e}_3 \times R_3(-\psi)R_1(-\theta)R_3(-\varphi)\mathsf{e}_3 \\
&= Mg\mu \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} \cos\psi & \sin\psi 0 & \\ \sin\psi & \cos\psi 0 & \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\
&= Mg\mu \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} \cos\psi & \sin\psi 0 & \\ \sin\psi & \cos\psi 0 & \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \sin\theta \\ \cos\theta \end{bmatrix} \\
&= Mg\mu \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} \sin\psi\sin\theta \\ \cos\psi\sin\theta \\ \cos\theta \end{bmatrix} \\
&= Mg\mu \begin{bmatrix} -\cos\psi\sin\theta \\ \sin\psi\sin\theta \\ 0 \end{bmatrix}
\end{aligned}$$

and obtain

$$I_1\dot{\omega}_1 - (I_2 - I_3)\omega_2\omega_3 = -Mg\mu\cos\psi\sin\theta$$

$$
\begin{aligned}
I_2\dot{\omega}_2 - (I_3 - I_1)\omega_3\omega_2 &= Mg\mu \sin\psi \sin\theta \\
I_3\dot{\omega}_3 - (I_1 - I_2)\omega_1\omega_2 &= 0
\end{aligned}
$$

Remembering that $I_1 = I_2$ we have

$$
\begin{aligned}
I_1\dot{\omega}_1 - (I_1 - I_3)J\omega_2 &= -Mg\cos\psi \sin\theta & (2.1) \\
I_1\dot{\omega}_2 + \frac{(I_1 - I_3)J}{I_3}\omega_1 &= Mg\sin\psi \sin\theta & (2.2) \\
I_3(\dot{\varphi}\cos\theta + \dot{\psi}) &= J = constant & (2.3)
\end{aligned}
$$

Expanding (2.1) and (2.2)

$$
\begin{aligned}
I_1 \quad & (\ddot{\varphi}\sin\theta\sin\psi + \dot{\varphi}\cos\theta\sin\psi\dot{\theta} + \dot{\varphi}\sin\theta\cos\psi\dot{\psi} + \ddot{\theta}\cos\psi - \dot{\theta}\sin\psi\dot{\psi}) \\
- \quad & (I_1 - I_3)\frac{J}{I_3}(\dot{\varphi}\sin\theta\cos\psi + \dot{\theta}\sin\psi) = -Mg\mu\cos\psi\sin\theta \qquad (2.4)
\end{aligned}
$$

and

$$
\begin{aligned}
I_1 \quad & (\ddot{\varphi}\sin\theta\cos\psi + \dot{\varphi}\cos\theta\cos\psi\dot{\theta} - \dot{\varphi}\sin\theta\sin\psi\dot{\psi} - \ddot{\theta}\sin\psi - \dot{\theta}\cos\psi\dot{\psi}) \\
+ \quad & (I_1 - I_3)\frac{J}{I_3}(\dot{\varphi}\sin\theta\sin\psi - \dot{\theta}\cos\psi) = Mg\mu\sin\psi\sin\theta \qquad (2.5)
\end{aligned}
$$

By multiplying (2.4) with $\cos\psi$ and (2.5) with $\sin\psi$ and subtracting we get

$$
I_1(\dot{\varphi}\dot{\psi}\sin\theta + \ddot{\theta}) - \frac{(I_1 - I_3)J}{I_3}(\dot{\varphi}\sin\theta) = -Mg\mu\sin\theta
$$

which gives upon substitution

$$
\dot{\psi} = \frac{J}{I_3} - \dot{\varphi}\cos\theta
$$

$$
I_1\ddot{\theta} - I_1\sin\theta\cos\theta\dot{\varphi}^2 + J\sin\theta\dot{\varphi} + Mg\mu\sin\theta = 0 \qquad (2.6)
$$

By multiplying (2.4) with $\sin\psi$, (2.5) with $\cos\psi$ and adding we obtain

$$
I_1(\sin\theta\ddot{\varphi} + 2\cos\theta\dot{\varphi}\dot{\theta}) - J\dot{\theta} = 0
$$

which gives upon multiplication with $\sin\theta$

$$
I_1(\sin^2\theta\ddot{\varphi} + 2 sin\theta cos\theta\dot{\varphi}\dot{\theta}) - J\sin\theta\dot{\theta} = 0
$$

which can be written in the form

$$\frac{d}{dt}\{I_1 \sin^2\theta\dot{\varphi} + J\cos\theta\} = 0 \tag{2.7}$$

Observe that the equations (2.6) and (2.7) are the same as (1.5) and (1.6) in Chapter 1.

CHAPTER 3

INSTANCES OF THE HAMILTONIAN APPROACH

In this and the following sections we shall denote vectors as

$$\mathsf{x} = \begin{bmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{bmatrix} \in \mathbb{R}^n \approxeq \mathbb{R}^{n\times 1},$$

covectors

$$\mathsf{y} = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} \in \mathbb{R}^n \approxeq \mathbb{R}^{1\times n}$$

and make use of the inner product $\langle \cdot, \cdot \rangle$, and cross-product in which we allow indiscriminate apperance of vector and covector arguments.

A large body of experience suggests that the proper mathematical setting for classical mechanics (among others) is a symplectic manifold which is an ordered pair (M, ω) where M is a manifold and ω is a symplectic form on M. On a sympectic manifold to each scalar field f a vector field X_f on M is associated which satisfies

$$df(Y) = \omega(\mathsf{X}_f, Y)$$

for every vector field Y on M.

A classical mechanics problem is formulated in many cases as a system of differential equations which is equivalent to the integration of a vector field of the form X_H on a symplectic manifold (M, ω) for some smooth $H : M \to \mathbb{R}$. The triple (M, ω, H) completely defines the problem, $H : M \to \mathbb{R}$ is referred to as the Hamiltonian and turns out to be constant along lines of the integral flow generated by X_H.

Example 3.1: Consider the cotangent bundle $M = T^*\mathbb{R}^n \approxeq \mathbb{R}^{2n}$ in which we fix each point by its coordinates $(x^i, y_i)_{1\leq i\leq n}$ where x^is refer to space components, y_is referring to covector components. Let $\omega = dx^i \wedge dy_i$ be the canonical symplectic form.

It can be routinely checked that

$$\mathsf{X}_F = \frac{\partial F}{\partial y_i}\frac{\partial}{\partial x^i} - \frac{\partial F}{\partial x^i}\frac{\partial}{\partial y_i}$$

for any scalar field $F \in C^\infty(\mathbb{R}^n)$. In particular if $H(\mathsf{x},\mathsf{y}) = \frac{1}{2} < \mathsf{y},\mathsf{y} >$, we obtain

$$\mathsf{X}_H = y_i\frac{\partial}{\partial x^i}$$

Thus the equations of motion read

$$\frac{d\mathsf{x}}{dt} = \mathsf{y} \qquad \frac{d\mathsf{y}}{dt} = 0$$

or

$$\frac{d^2\mathsf{x}}{dt} = 0$$

These are clearly the equations for motion of a free particle.

Example 3.2 : Let again $M = T^*\mathbb{R}^3 \approxeq \mathbb{R}^6$, $\omega = dx^i \wedge dy_i$ and let

$$H(\mathsf{x},\mathsf{y}) = \frac{1}{2} < \mathsf{y},\mathsf{y} > +\gamma x^3$$

By a similar calculation as in (**Example 3.1**) we can show that

$$\mathsf{X}_H = -y_i\frac{\partial}{\partial x^i} - \gamma\frac{\partial}{\partial y_3}$$

and the equations of motion read

$$\frac{d\mathsf{x}}{dt} = \mathsf{y} \qquad \frac{d\mathsf{y}}{dt} = -\gamma\mathsf{e}_3$$

or equivalently,

$$\frac{d^2\mathsf{x}}{dt} = -\gamma\mathsf{e}_3$$

This is obviously the equation for the motion of a free particle under constant uniform gravitational field.

Example 3.3: Let, once again, $M = T^*\mathbb{R}^3 \approxeq \mathbb{R}^6$, $\omega = dx^i \wedge dy_i$ and introduce a modified symplectic form Ω defined by

$$\Omega = \omega + \frac{\mu}{\|\mathsf{x}\|^3}\ \epsilon_{ijk}\ x^i\ dx^j \otimes dx^k.$$

For an arbitrary scalar field $F \in C^\infty(M)$ it can be routinely checked that

$$\mathsf{X}_F = \frac{\partial F}{\partial y_i}\frac{\partial}{\partial x^i} + \left(-\frac{\partial F}{\partial x^i} + \frac{\mu}{\|\mathsf{x}\|^3}\ \epsilon_{ijk}\ x^j\ \frac{\partial F}{\partial y_k}\right)\frac{\partial}{\partial y_i}$$

In particular, if

$$H(\mathsf{x},\mathsf{y}) = \frac{1}{2}\langle \mathsf{y},\mathsf{y}\rangle + \gamma\langle \mathsf{x},\mathsf{e}_3\rangle$$

then

$$\mathsf{X}_H = y_i\frac{\partial}{\partial x^i} + (-\gamma\delta_{i3} + \frac{\mu}{\|\mathsf{x}\|^3}\epsilon_{ijk}\ x^i y_k)\frac{\partial}{\partial y_i}$$

and the equations of motion read

$$\frac{d\mathsf{x}}{dt} = \mathsf{y}, \qquad \frac{d\mathsf{y}}{dt} = -\gamma\mathsf{e}_3 + \frac{\mu}{\|\mathsf{x}\|^3}\mathsf{x}\times\mathsf{y}$$

These equations clearly govern the motion of a massive, charged free particle moving under the action of a uniform homogeneous gravitational field and a magnetic monopole placed at the origin.

An important development emanating from Hamiltonian mechanics is the concept of Poisson brackets.

Definition 3.1: Quite generally a *Poisson algebra* is a commutative algebra $\mathfrak{A}$ over $\mathbb{R}$ with a bilinear binary operation $\{\cdot,\cdot\} : \mathfrak{A}\times\mathfrak{A} \to \mathfrak{A}$ such that

1. {f, g}= -{g, f}

2. {f, { g, h}}+ {g, {h, f }}+{h, {f, g}}= 0

3. {f, gh }= {f, g} h+ g {f, h}

for any $f, g \in \mathfrak{A}$.

In connection with a symplectic manifold (M, ω) there exists a natural Poisson algebra structure on $C^\infty(M)$. The Poisson bracket $\{f, g\}$ of $f, g \in C^\infty(M)$ is defined by

$$\{f, g\} = \omega(\mathsf{X_f}, \mathsf{X_g}).$$

It is important to notice that the Poisson bracket defined in the above fashion depends on the choice of the symplectic form.

Example 3.4: For the standard situation $M = T^*\mathbb{R}^n \approxeq \mathbb{R}^{2n}$, $\omega = dx^i \wedge dy_i$ the Poisson bracket becomes

$$\begin{aligned}\{f, g\} &= dx^i \wedge dy_i(\frac{\partial f}{\partial x^q}\frac{\partial}{\partial y_q} - \frac{\partial f}{\partial y_q}\frac{\partial}{\partial x^q}, \frac{\partial g}{\partial x^r}\frac{\partial}{\partial y_r} - \frac{\partial g}{\partial y_r}\frac{\partial}{\partial x^r}) \\ &= \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial y_i} - \frac{\partial f}{\partial y_i}\frac{\partial g}{\partial x^i} \\ &= \langle \frac{\partial f}{\partial \mathsf{x}}, \frac{\partial g}{\partial \mathsf{y}} \rangle - \langle \frac{\partial f}{\partial \mathsf{y}}, \frac{\partial g}{\partial \mathsf{x}} \rangle\end{aligned}$$

Example 3.5: For the symplectic manifold $(T^*\mathbb{R}^3, \Omega)$, the Poisson bracket takes the form

$$\{f, g\} = \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial y_i} - \frac{\partial f}{\partial y_i}\frac{\partial g}{\partial x^i} - \frac{\mu}{\|\mathsf{x}^3\|}\epsilon_{ijk} x^i \frac{\partial f}{\partial y_j}\frac{\partial g}{\partial y_k}$$

Example 3.6: Consider a symplectic manifold (M, ω) quite generally. In the presence of a chart $x = (x^i)_{1 \le i \le n}$ on M, let

$$\omega|_{dom(x)} = \omega_{ij} dx^i \otimes dx^j$$

and let $[\omega^{ij}]_{1 \le i,j \le n}$ be the inverse of the matrix $[\omega_{ij}]_{1 \le i,j \le n}$. For any $F \in C^\infty(M)$, as

$$\omega(X_F, \frac{\partial}{\partial x^j})|_{dom(x)} = \frac{\partial F}{\partial x^j}$$

we conclude that

$$X_F|_{dom(x)} = \omega^{ij}\frac{\partial F}{\partial x^j}\frac{\partial}{\partial x^i}$$

Notice in particular, that

$$X_{x^q} = \omega^{iq}\frac{\partial}{\partial x^i}$$

and hence

$$\begin{aligned}\{x^p, x^q\} &= \omega_{ij}\omega^{ip}\omega^{jq} \\ &= \omega_{pq}\end{aligned}$$

The importance of this approach lies in the following observation due to Poisson: In the presence of a Hamiltonian H, one considers the "dynamical system", i.e the flow φ_t(or local flow) on M, generated by X_H. For any scalar field $f \in C^\infty(M)$, the development of $f(t) = f(\varphi_t(m))$ for any $m \in M$ in "time" is described by

$$\frac{df}{dt} = X_H f = df(X_H) = \omega(X_f, X_H)$$

hence

$$\frac{df}{dt} = \{f, H\}.$$

This shows that the Poisson algebra structure on $C^\infty(M)$ alone is sufficient for a full description of the time development of any "observable" $f \in C^\infty(M)$. It is an interesting philosophical consideration to notice that as M is a manifold and each point $m \in M$ is fully described by choosing a local coordinate system $x = (x^i)_{1\le i\le n}$ on M with $m \in dom(x)$ and assigning m the numbers $(x^i(m))_{1\le i\le n}$, one can write equations of motion of a Hamiltonin system in the form

$$\frac{dx^i}{dt} = \{x^i, H\}.$$

Of course one has to bear in mind the slight difficulty that arises in view of the fact that the coordinates x^i are "local" observable only.

Thus, there is no loss of dynamical information in passing from a symplectic system(M,ω) with a distinguished observable H to the Poisson algabra

$(C(M), \{\cdot\})$ with distinguished element of $H \in C^\infty(M)$. This theme is to be developed a little further.

For which reason there is much research conducted on Poisson algebras as pure algebraic entities on the one hand and the so-called Poisson manifolds. A Poisson manifold is a manifold M in which $C^\infty(M)$ as a commutative algebra over $\mathbb{R}$ admits a bilinear binary operation $\{\cdot,\cdot\}$ obeying the properties 1,2 and 3 above.

As we have seen above, symplectic manifolds have natural structures as Poisson manifolds. However, there are Poisson manifolds which do not arise from symplectic manifolds. Indeed, symplectic manifolds are even dimensional whereas it is possible to construct odd dimensional Poisson manifolds.

Example 3.7: $\mathbb{R}^3$ has a Poisson manifold structure. Take and fix an arbitrary $F \in C^\infty(\mathbb{R}^3)$. It can be checked that $\{\cdot,\cdot\}$ on $C^\infty(\mathbb{R}^3)$ defined by

$$\begin{aligned}\{f,g\} &= \epsilon_{ijk}\frac{\partial f}{\partial x^j}\frac{\partial g}{\partial x^k}\frac{\partial F}{\partial x^i} \\ &= (\nabla f \times \nabla g)\cdot \nabla F \\ &= det(\nabla f, \nabla g \nabla F)\end{aligned}$$

gives rise to a Poisson algebra.

Concerning the Poisson formalism in manifolds one final remark is due : At first sight it seems impossible to produce a finite table which gives a complete description of $\{\cdot,\cdot\}$ on $C^\infty(M)$ (which is infinite dimensional) as in the case of finite dimensional Lie algebras. Upon closer inspection it will be seen to be possible if the manifold can be covered by finitely many charts(clearly finite). In this case a multiplication table for the components of charts will be sufficient to determine everything else, since given $f, g \in C^\infty(M)$ and a chart $x = (x^i)_{1\leq i\leq n}$ we have

$$\begin{aligned}\{f,g\}|_{dom(x)} &= \omega(X_f, X_g)|_{dom(x)} \\ &= \omega_{ij}\omega^{ip}\frac{\partial f}{\partial x^p}\omega^{jq}\frac{\partial g}{\partial x^q} \\ &= -\frac{\partial f}{\partial x^j}\frac{\partial g}{\partial x^q}\omega^{jp}\end{aligned}$$

$$
\begin{aligned}
&= \frac{\partial f}{\partial x^j}\frac{\partial g}{\partial x^q}\omega^{qj} \\
&= \frac{\partial f}{\partial x^j}\frac{\partial g}{\partial x^q}\{x^j, x^q\} \\
&= \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial x^j}\{x^i, x^j\}
\end{aligned}
$$

CHAPTER 4

HAMILTONIAN DYNAMICS ON SUBMANIFOLDS VIA THE DIRAC PRESCRIPTION

Given a Hamiltonian system (M, ω, H) it can be important to restrict the problem to a submanifold N of M. Of course N has to be even dimensional. This condition is often automatically fulfilled since typically one has $M = T^*Q$ for some manifold Q and $N = T^*V$ for some submanifold V of Q. However even if this obvious dimension condition is fulfilled the restriction of the symplectic form to the submanifold can easily fail to be non-degenerate.

Example 4.1: Let $M = \mathbb{R}^4$ with the symplectic form $\omega = dx \wedge dy + du \wedge dv$. Consider the submanifold $A = \{(x, x, u, u) | x, u \in \mathbb{R}\}$ of $\mathbb{R}^4$. We can easily show that $\omega(\mathsf{s}, \mathsf{t}) = 0$ for typical tangent vectors s, t to A.

Of course it is possible to deal with each special case separately. However, the difficulty is alleviated for a large class of submanifols by the following theorem usually ascribed to P. Dirac ([Dir], see [C-B 1] for a modern mathematical treatment):

Theorem 4.1: Let (M, Ω) be a symplectic manifold and N be a submanifold of M where $N = F^{-1}(q)$ for some regular value $q \in M$ of $F = (F_1, F_2, \ldots, F_r) : M \to \mathbb{R}^r$. $\Omega|_N$ is a symplectic form on N if the matrix

$$\Big[\ \{F_i, F_j\} \ \Big]_{1 \le r, j \le r}$$

is non- singular at each point of N.

Proof: Appendix.

Example 4.2: Let $M = T^*(\mathbb{R}^3 - \{(0, 0, 0)\})$ with the canonical symplectic form $\omega = dx^i \wedge dy_i$ and a submanifold $N = T^*S^2$ defined by

$$F_1(\mathsf{x}, \mathsf{y}) = \langle \mathsf{x}, \mathsf{x} \rangle - 1 = 0$$
$$F_2(\mathsf{x}, \mathsf{y}) = \langle \mathsf{x}, \mathsf{y} \rangle = 0$$

Then

$$X_{F_1} = -2x^i \frac{\partial}{\partial y_i}, \qquad X_{F_2} = x^i \frac{\partial}{\partial x^i} - y^i \frac{\partial}{\partial y_i}$$

From $\{F_i, F_j\} = dx^i \wedge dy_i(X_{F_i}, X_{Fj})$ for $i = 1, 2$ we can find the matrix

$$\begin{bmatrix} 0 & \{F_1, F_2\} \\ \{F_2, F_1\} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$$

is non- singular at each point point of T^*S^2. Then $\Omega|_N$ is a sympectic form on N.

Apart from the above mentioned difficulty which can be taken care of by the above theorem there are other inherent difficulties in dealing with submanifolds. For instance as a bonus for working with the Poisson formalism one might hope to deal with such difficulties by noticing that each element of $C^\infty(N)$ the restriction to N of some element of $C^\infty(M)$ and expecting that the Poisson bracket of $f|_N$ and $g|_N$ on N is the restriction of $\{f, g\}$ to N for all $f, g \in C^\infty(M)$. However, this is not true. To be precise, it is in general not true that for any scalar fields f, g on M

$$\{f|_N, g|_N\}_N \neq \{f, g\}|_N$$

where $\{\cdot, \cdot\}_N$ stands for the Poisson bracket inducted by the symplectic form $\omega|_N$ on N.

Example 4.3: Given $M = \mathbb{R}^4$ and $\omega = dx \wedge dy + dz \wedge dt$. Let $N = \mathbb{R}^2$. Clearly $i_*(\omega) = dx \wedge dy$. Define

$$F(x, y, z, t) = f(x, y, z, t) + z$$
$$G(x, y, z, t) = g(x, y, z, t) + t$$

for some scalar functions f and g.

A simple calculation shows that

$$\{F|_{\mathbb{R}^2}, G|_{\mathbb{R}^2}\}_{\mathbb{R}^2} = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}$$

On the other hand

$$\{F, G\}|_{\mathbb{R}^2} = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x} + 1$$

Therefore

$$\{F|_{\mathbb{R}^2}, G|_{\mathbb{R}^2}\}_{\mathbb{R}^2} \neq \{F, G\}_{\mathbb{R}^4}|_{\mathbb{R}^2}$$

Once again it is possible to save a considerable piece of the theory by means of the following theorem due to P. M. Dirac([Dir]):

Theorem 4.2: For any $f, g \in C^\infty(M)$

$$\{f|_N, g|_N\}_N = \{f, g\}^*|_N$$

where bracket $\{\cdot, \cdot\}^*$ on $C^\infty(M)$ is defined by

$$\{f, g\}^* = \{f, g\} - \{f, F_i\}\Psi^{ij}\{F_j, g\}$$

with the matrix

$$\left[\ \Psi^{ij}\ \right]_{1\leq i,j\leq r}$$

being the inverse of the matrix

$$\left[\ \{F_i, F_j\}\ \right]_{1\leq i,j\leq r}$$

Proof: Appendix.

In our case, $M = T^*(\mathbb{R}^3 - \{(0,0,0)\})$ and a submanifold $N = T^*S^2$ is defined by

$$\begin{aligned} F_1(\mathsf{x}, \mathsf{y}) &= \langle \mathsf{x}, \mathsf{x} \rangle - 1 = 0 \\ F_2(\mathsf{x}, \mathsf{y}) &= \langle \mathsf{x}, \mathsf{y} \rangle = 0. \end{aligned}$$

T^*S^2 is a symplectic manifold of $T^*(\mathbb{R}^3 - \{(0,0,0)\})$ with the symplectic form $\widetilde{\Omega} = i_*(\Omega)$ where i is the inclusion map from $N = {F_1}^{-1}(0) \cap {F_2}^{-1}(0)$ to M since the matrix

$$\begin{bmatrix} 0 & \{F_1, F_2\} \\ \{F_2, F_1\} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$$

is non-singular at each point point of T^*S^2. Hence,

$$\Psi^{ij} = \begin{bmatrix} 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$$

Thus we find that

$$\{f, g\}^* = \{f, g\} + \frac{1}{2}\{f, F_1\}\{F_2, g\} - \frac{1}{2}\{f, F_2\}\{F_1, g\}$$

More explicitly,

$$\begin{aligned} \{f, g\}^* &= \{f, g\} + \langle \frac{\partial f}{\partial \mathsf{y}}, \mathsf{x}\rangle [\langle \frac{\partial g}{\partial \mathsf{x}}, \mathsf{x}\rangle - \langle \frac{\partial g}{\partial \mathsf{y}}, \mathsf{y}\rangle] \\ &- \langle \frac{\partial g}{\partial \mathsf{y}}, \mathsf{y}\rangle [\langle \frac{\partial f}{\partial \mathsf{x}}, \mathsf{x}\rangle - \langle \frac{\partial f}{\partial \mathsf{y}}, \mathsf{y}\rangle] \end{aligned}$$

The magnetic spherical pendulum is the Hamiltonian system $(T^*S^2, \Omega|_{T^*S^2}, H|_{T^*S^2})$ where

$$\begin{aligned} H: \; T^*\mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (\mathsf{x}, \mathsf{y}) &\longmapsto \frac{1}{2}\langle \mathsf{y}, \mathsf{y}\rangle + \gamma\langle \mathsf{x}, \mathsf{e}_3\rangle \end{aligned}$$

Instead of using $H|_{T^*S^2}$ we can use Dirac's prescription to compute equations of motion by using the new bracket $\{\cdot, \cdot\}^*$.
In this case the equations of motion become

$$\frac{dx^i}{dt} = \{x^i, H\}^*, \qquad \frac{dy_i}{dt} = \{y_i, H\}^*$$

Then

$$\begin{aligned} \{f, F_1\} &= -\frac{\partial f}{\partial y_i}\, 2x^i = -2\langle \frac{\partial f}{\partial \mathsf{y}}, \mathsf{x}\rangle \\ \{f, F_2\} &= \frac{\partial f}{\partial x^i}\, x^i - \frac{\partial f}{\partial y_i}\, y_i - \frac{\mu}{\|\mathsf{x}\|^3}\epsilon_{pjk} x^p \frac{\partial f}{\partial y_j}\, x^k = \langle \frac{\partial f}{\partial \mathsf{x}}, \mathsf{x}\rangle - \langle \frac{\partial f}{\partial \mathsf{y}}, \mathsf{y}\rangle \end{aligned}$$

Then

$$\begin{aligned}
\frac{dx^i}{dt} &= \{x^i, H\}^* = y_i \\
\frac{dy_i}{dt} &= \{y_i, H\}^* = \frac{\partial y_i}{\partial x^p} - \frac{\partial y_i}{\partial y_p}\gamma\delta_{3p} - \frac{\mu}{\|\mathsf{x}\|^3}\epsilon_{pjk}x^p\frac{\partial y_i}{\partial y_j}\frac{\partial H}{\partial y_k}
\end{aligned}$$

Finally, it follows from the discussion above that the equations of motion of the magnetic spherical pendulum are

$$\frac{d\mathsf{x}}{dt} = \mathsf{y} \qquad \frac{d\mathsf{y}}{dt} = -\gamma\mathsf{e}_3 + [< \mathsf{e}_3, \mathsf{x} > \gamma - < \mathsf{y}, \mathsf{y} >]\mathsf{x} + \frac{\mu}{\|\mathsf{x}\|^3}]\mathsf{x} \times \mathsf{y}$$

by using $\langle \mathsf{x}, \mathsf{y} \rangle = 0$ on T^*S^2.

CHAPTER 5

GENERAL REMARKS ON REDUCTION IN CLASSICAL MECHANICS

In the framework of mechanics the word *reduction* loosely refers to the process of obtaining a lower dimensional phase space by taking a suitable quotient in the presence of a symmetry. Ideally the phase space is a manifold M, the symmetry is described as the action of a Lie group G on M preserving everything mechanically relevant and the quotient space M/G is a manifold, too. This ideal situation fails to arise in many important cases. Indeed, despite numerous attempts at clarifying it, the word reduction seems still to connote little more than a gentlemen' s agreement at present.

One should perhaps start with the most basic and best known classical instance:

Example 5.1: Consider the planar motion of a material particle under the influence of a central conservative force field. In this case F will be

$$\begin{aligned} \mathsf{F} &= -grad\, V \\ &= -\frac{\partial V}{\partial x}\frac{\partial}{\partial x} - \frac{\partial V}{\partial y}\frac{\partial}{\partial y} \end{aligned}$$

for some potential function $V = V(x, y)$. The condition that F is always directed towards $(0, 0)$ can be expressed by

$$y\frac{\partial V}{\partial x} - x\frac{\partial V}{\partial y} = 0$$

which also shows that, of the usual polar coordinates (r, θ) on $\mathbb{R}^2-(\{0\}\times[0,\infty))$

V is independent of θ : Indeed

$$\begin{aligned}\frac{\partial V}{\partial \theta} &= \frac{\partial V}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial V}{\partial y}\frac{\partial y}{\partial \theta} \\ &= \frac{\partial V}{\partial x}(-r\sin\theta) + \frac{\partial V}{\partial y}(r\cos\theta) \\ &= -y\frac{\partial V}{\partial x} + x\frac{\partial V}{\partial y} = 0\end{aligned}$$

Here we have

$$\mathsf{x} = r\mathsf{e}_\mathsf{r}$$

where

$$\mathsf{e}_r = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

hence

$$\frac{d\mathsf{x}}{dt} = \dot{r}\mathsf{e}_\mathsf{r} + r\mathsf{e}_\theta\dot{\theta}$$

where

$$\mathsf{e}_\theta = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

Thus

$$\frac{d^2\mathsf{x}}{dt^2} = \ddot{r}\mathsf{e}_r + \dot{r}\mathsf{e}_\theta\dot{\theta} + \dot{r}\mathsf{e}_\theta\dot{\theta} + r\mathsf{e}_\theta\ddot{\theta} - r\dot{\theta}^2\mathsf{e}_r$$

or equivalently

$$\frac{d^2\mathsf{x}}{dt^2} = (\ddot{r} - r\dot{\theta}^2)\mathsf{e}_r + (2\dot{r}\dot{\theta} + r\ddot{\theta})\mathsf{e}_\theta$$

On the other hand

$$\mathsf{F} = -\frac{\partial V}{\partial \mathsf{r}} = -\frac{\partial V}{\partial x}\mathsf{e}_1 + \frac{\partial V}{\partial y}\mathsf{e}_2$$

where

$$\begin{aligned}\mathsf{e}_1 &= \cos\theta\mathsf{e}_r - \sin\theta\mathsf{e}_\theta \\ \mathsf{e}_2 &= \sin\theta\mathsf{e}_r + \cos\theta\mathsf{e}_\theta\end{aligned}$$

Therefore

$$\begin{aligned}\mathsf{F} &= (-\frac{\partial V}{\partial x}\cos\theta + \frac{\partial V}{\partial y}\sin\theta)\mathsf{e}_r + (\frac{\partial V}{\partial x}\sin\theta + \frac{\partial V}{\partial y}\cos\theta)\mathsf{e}_\theta \\ &= -\frac{\partial V}{\partial r}\mathsf{e}_r\end{aligned}$$

Consequently the equations of motion read

$$\begin{aligned} 2\dot{r}\dot{\theta} + r\ddot{\theta} &= 0 \\ m(\ddot{r} - r\dot{\theta}^2) &= -\frac{\partial V}{\partial r} \end{aligned}$$

Thus $mr^2\dot{\theta}$ is a constant which we denote by l and write

$$mr\dot{\theta}^2 = \frac{l^2}{mr^3}$$

and the problem is reduced to the 1-dimensional problem

$$m\frac{d^2r}{dt^2} = -\frac{\partial \widetilde{V}}{\partial r}$$

where

$$\widetilde{V}(r) = V(r) + \frac{l^2}{2mr^2}$$

In classical texts the new potential function $\widetilde{V} = \widetilde{V}(r)$ is called the "effective potential".

Given a Lie group G and a manifold M, a (left) *action* of G on M is a smooth map

$$\sigma : G \times M \longrightarrow M$$

such that

$$\begin{aligned} \sigma(g_2, \sigma(g_1, m)) &= \sigma(g_2 g_1, m) \\ \sigma(e, m) &= m \end{aligned}$$

for all $m \in M$, $g_1, g_2 \in G$ where $e \in G$ is the neutral element of G. Once it is clear which group action is being considered and there is no rise of confusion we shall write gm instead of $\sigma(g, m)$.

Of course flows constitute the most important class of group actions with $G = \mathbb{R}$.

Remember that given topological spaces $X,\ Y$ a map $f : X \rightarrow Y$ is called *proper* if the inverse image of every compact subset of Y under f is compact in X. An action of the Lie group G on M is called proper if the map

$$G \times M \longrightarrow M \times M$$

sending (g, m) into (gm, m) is proper. Notice that all actions of a compact Lie group are proper.

For each $m \in M$, the set

$$G_m = \{g \in G | gm = m\}$$

constitutes a subgroup of G which is called the *isotropy group* of the action at $m \in M$. An action is called *free* if G_m is trivial for each $m \in M$.

Let $\mathfrak{g}$ denote the Lie algebra of G consisting of the left invariant vector fields on G. For every $A \in \mathfrak{g}$ there exist a vector field A_* on M such that $(t, m) \rightarrow exp(tA)$ is the flow on M induced by A_*.

Example 5.2: Consider the Lie group $G = S^1 = SO(2) \approxeq \mathbb{R}/2\pi\mathbb{Z}$ with its Lie algebra

$$\mathfrak{g} = so(2) \approxeq \mathbb{R} = \langle \frac{\partial}{\partial\theta} \rangle$$

and its action on $\mathbb{R}^2$ defined by

$$[\theta](x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$$

where $[\theta] \in S^1 \approxeq \mathbb{R}/2\pi\mathbb{Z}$ denotes the residue class of $2\pi\mathbb{Z}$ in $\mathbb{R}$ containing θ. Note that this is the action which characterises the classical problem of "central force fields" discussed at the beginning.

As $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ is compact this action is automatically proper. However it fails to be free since

$$G_{(x,y)} = [0] \in S^1$$

for all $(x, y) \neq (0,0) \in \mathbb{R}^2$ but $G_{(0,0)} = S^1$.

We have on S^1

$$exp(t\frac{\partial}{\partial\theta}) = [t]$$

and hence

$$\begin{aligned}(\frac{\partial}{\partial\theta})_*|_{(x,y)} &= \frac{d}{dt}|_{t=0}(x\cos t - y\sin t, x\sin t + y\cos t)\\ &= -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}|_{(x,y)}\end{aligned}$$

Given an action of G on M the *orbit* through m is the set

$$Orb(m) = \{gm|g \in G\}.$$

Clearly M is the disjoint union of orbits.

Let M/G denote the set of orbits, let

$$\pi : M \longrightarrow M/G$$

be the map sending $m \in M$ into $Orb(m)$. We shall consider M/G with the quotient topology induced on it by π.
It is known that M/G has a natural structure as a smooth manifold such that π is a smooth map if the action of G is free and proper. In the above example the action of S^1 on $\mathbb{R}^2$ is compact but not free. Indeed, $\mathbb{R}^2/S^1$ is homeomorphic to $[0, \infty)$ and therefore it is not a manifold.

Definition 5.1: Action of G on a symplectic manifold (M, ω) is called *Hamiltonian* if for each $A \in \mathfrak{g}$ there exist a scalar field $J_A \in C(M)$ with $A_* = \mathsf{X}_{J_A}$. The map $J : M \to \mathfrak{g}^*$ is called the *momentum map* of the Hamiltonian action. Notice that J_A depends linearly on $A \in \mathfrak{g}$.

Clearly, so does $J_A(m)$ depends linearly on $A \in \mathfrak{g}$. As a consequence, it is more convenient to think of J as a map from M to $\mathfrak{g}^*$(the dual of $\mathfrak{g}$, as a vector space, that is, with its Lie algebra structure forgotten) assigning to each $m \in M$ the linear map sending $A \in \mathfrak{g}$ to $J_A(m) \in \mathbb{R}$, clearly an elementof $\mathfrak{g}^*$.

Example 5.3: ([L-M-S]) Consider the $S^1 = SO(2)$ action on $\mathbb{R}^2$ extended to $T^*\mathbb{R}^2$ with the canonical symplectic form $\omega = dx^1 \wedge dy_1 + dx^2 \wedge dy_2$ defined by

$$\begin{bmatrix} x^1 \\ x^2 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x^1 \cos t - x^2 \sin t \\ x^1 \sin t + x^2 \cos t \\ y_1 \cos t - y_2 \sin t \\ y_1 \sin t - y_2 \cos t \end{bmatrix}$$

It can be checked that this action is Hamiltonian and admits

$$J = x^1 y_2 - x^2 y_1$$

as momentum map. If one wishes to be pedantic, one might insist on remembering $G = SO(2), \mathfrak{g} = so(2)$ and considers J to be actually a map from M into $so(2)^* = \langle d\theta \rangle$ and put

$$J = (x^1 y_2 - x^2 y_1) d\theta$$

We do not. Indeed, it can be routinly checked that

$$X_{x^1 y_2 - x^2 y_1} = -x^2 \frac{\partial}{\partial x^1} + x^1 \frac{\partial}{\partial x^2} - y_2 \frac{\partial}{\partial y_1} + y_2 \frac{\partial}{\partial y_1}$$

Clearly, each set of the form $J^{-1}(a)$ is invariant under the action of $SO(2) \approxeq \mathbb{R}/2\pi\mathbb{Z}$. In this particular case

$$(-x^2 \frac{\partial}{\partial x^1} + x^1 \frac{\partial}{\partial x^2} - y_2 \frac{\partial}{\partial y_1} + y_2 \frac{\partial}{\partial y_2})(x^1 y_2 - x^2 y_1) = 0$$

More generally

$$\frac{d}{dt} J = X_J J = \{J, J\} = 0$$

When $a \in \mathbb{R}$ is a regular value of J then $J^{-1}(a)$ is a 3 dimensional submanifold of $T^*\mathbb{R}^2 \approxeq \mathbb{R}^4$ which is invariant under the action of $S^1 \approxeq SO(2)$. In such a situation we may consider the problem to have been effectively reduced to a lower dimensional problem. Of course $J^{-1}(a)$ is odd dimensional and the restriction of ω to $J^{-1}(a)$ is not non-degenerate. However this difficulty can be surmounted by taking quotient under $S^1 \approxeq SO(2)$ and noticing that the nondegeneracy of ω is lifted by this means. This process will be revisited below.

$0 \in \mathbb{R}$, however, is *not* a regular value of J. In this case we observe that

$$J^{-1}(0) = K_0 \cup K_1$$

where

$$\begin{aligned} K_0 &= \{(0,0,0,0)\} \\ K_1 &= \{(x^1, x^2, y_1, y_2) \in T^*\mathbb{R}^2 | x^1 y_2 - x^2 y_1 = 0, (x^1, x^2, y_1, y_2) \neq (0,0,0,0)\} \end{aligned}$$

K_1 is a submanifold. Moreover, the action of $S^1 \approxeq SO(2)$ on K_1 is free and (trivially) proper. Therefore $K_1/SO(2)$ is a two dimensional smooth manifold. An interesting and useful coincidence(or not?) is the following. The restriction of ω to K_1 is degenerate. But it can be checked that for every $m \in K_1$

$$(T_m K_1)^{\perp} = \{m \in T_m T^*\mathbb{R}^2 |\ \omega(u,v) = 0\ \forall\ v \in T_m K_1\} \subseteq T_m K_1$$

and indeed $(T_m K_1)^{\perp}$ is exactly the one dimensional subspace of $T_m K_1$ generated by the action of $S^1 \approxeq SO(2)$ on K_1. Consequently, the exterior form $\tilde{\omega} = \omega/SO(2)$ induced on K_1 is non-degenerate, that is $(K_1/SO(2), \tilde{\omega})$ is a smooth symplectic manifold. K_0 is a singleton on which $S^1 = SO(2)$ acts trivially. It is, however, to be noticed, that $J^{-1}(0) = K_0 \cup K_1$ is *not* a smooth manifold.

One can obtain a more concrete picture of $J^{-1}(0)$ by considering

$$\Lambda = \{(x^1, 0, y_1, 0) \in T^*\mathbb{R}^2\}$$

and noticing that the null-set $J^{-1}(0)$ of the momentum map J completely contains Λ which is a smooth submanifold of $T^*\mathbb{R}^2$ diffeomorphic to $\mathbb{R}^2$ and the S^1-orbit of any point $(\mathsf{x}, \mathsf{y}) \in J^{-1}(0)$ intersects Λ in exactly two points.

Moreover, a point (x, y) in $J^{-1}(0)$ if and only if x and y are collinear as vectors in $\mathbb{R}^2$.
Thus $J^{-1}(0)$ may be obtained by identifying (x, y) and $(-\mathsf{x}, -\mathsf{y})$ in $\mathbb{R}^2$.

There is yet another way of investigating the dynamics on $J^{-1}(0)/SO(2)$ rather than its structure.

Following polynomials in $\mathbb{R}[x^1, x^2, y_1, y_2]$ are invariant under the action of $S^1 = SO(2)$.

$$\begin{aligned} X &= (x^1)^2 + (x^2)^2 \\ Y &= (y_1)^2 + (y_2)^2 \\ Z &= x^1 y_1 + x^2 y_2 \\ U &= x^1 y_2 - x^2 y_1 \end{aligned}$$

It is known that each polynomial in $\mathbb{R}[x^1, x^2, y_1, y_2]$ which is invariant under the action of $S^1 = SO(2)$ can be expressed as a polynomial in X, Y, Z, U and this set is minimal.([Weyl])

With the map

$$\begin{aligned} \Theta : T^*\mathbb{R}^2 &\longrightarrow \mathbb{R}^4 \\ (x^1, x^2, y_1, y_2) &\longmapsto (X, Y, Z, U) \end{aligned}$$

$\Theta(J^{-1}(0)) \in \mathbb{R}^4$ may be identified with $\mathbb{R}^3 = \{(X, Y, Z, U) \in \mathbb{R}^4 |\ U = 0\}$.
Therefore the invariants satisfy the equations

$$\begin{aligned} Z^2 &= XY \\ X, Y &\geqslant 0 \\ U &= 0 \end{aligned}$$

This is a topological submanifold of $\mathbb{R}^3$ which is smooth except at $(0, 0, 0)$.

CHAPTER 6

SPHERICAL PENDULUM

In this chapter, we will construct the Poisson brackets on a suitably reduced phase space for the spherical pendulum by using the method invariant polynomials as presented in the previous chapter.

Consider the action of $S^1 = SO(2) \approxeq \mathbb{R}/2\pi\mathbb{Z}$ on $(T^*\mathbb{R}^3, \omega)$ defined by

$$\begin{aligned} SO(2) \times T^*\mathbb{R}^3 &\longrightarrow T^*\mathbb{R}^3 \\ ([t], (\mathsf{x}, \mathsf{y})) &\longmapsto (R_t\mathsf{x}, R_t\mathsf{y}) \end{aligned}$$

where

$$R_t = \begin{bmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note that this map leaves the subspace defined by $\{x^3 = y_3 = 0\}$ invariant.

The following polynomials are invariant under this action of S^1 and they generate the algebra of all S^1-invariant polynomials in $\mathbb{R}[x^1, x^2, y_1, y_2]$(weyl)

$$\begin{aligned} X &= x^3 \\ Y &= y_3 \\ Z &= (y_1)^2 + (y_2)^2 + (y_3)^2 \\ U &= x^1 y_1 + x^2 y_2 \\ V &= (x^1)^2 + (x^2)^2 \\ W &= x^1 y_2 - x^2 y_1 \end{aligned}$$

Define the Hilbert map σ for the S^1- action as

$$T^*\mathbb{R}^3 \approxeq \mathbb{R}^6 \longrightarrow \mathbb{R}^6$$

$$(x^1, x^2, x^3, y_1, y_2, y_3) \longmapsto (X, Y, Z, U, V, W)$$

Then (X, Y, Z, U, V, W) coordinates satisfy

$$\begin{aligned} U^2 + W^2 &= V(Z - Y^2) \\ Z &\geqslant 0 \\ V &\geqslant 0 \end{aligned}$$

The non-zero Poisson brackets of X, Y, Z, U, V, W are

$$\begin{aligned} \{X, Y\}_{\mathbb{R}^6} &= 1 \\ \{X, Z\}_{\mathbb{R}^6} &= 2Y \\ \{Z, U\}_{\mathbb{R}^6} &= -2(Z - Y^2) \\ \{Z, V\}_{\mathbb{R}^6} &= -4U \\ \{U, V\}_{\mathbb{R}^6} &= -2V \end{aligned}$$

Recall that submanifold T^*S^2 of $T^*\mathbb{R}^3$ is defined by the equations

$$\begin{aligned} F_1(\mathsf{x}, \mathsf{y}) &= \langle \mathsf{x}, \mathsf{x} \rangle - 1 = V + X^2 - 1 = 0 \\ F_2(\mathsf{x}, \mathsf{y}) &= \langle \mathsf{x}, \mathsf{y} \rangle = U + XY = 0 \end{aligned}$$

It can be routinly checked that T^*S^2 is invariant under the action of S^1. $\omega|T^*S^2$ is a symplectic form on T^*S^2 since

$$[\{F_i, F_j\}]_{1 \le i,j \le 2} = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$$

is non-singular at each point of T^*S^2.
Therefore the image of T^*S^2 under σ consists of points $(X, Y, Z, U, V, W) \in \mathbb{R}^6$ satisfying

$$\begin{aligned} U^2 + W^2 &= V(Z - Y^2) \\ V + X^2 &= 1 \\ U + XY &= 0 \\ Z &\geqslant 0 \\ V &\geqslant 0 \end{aligned}$$

It can be checked that the action of $S^1 \approxeq SO(2)$ on $T^*\mathbb{R}^3$ is Hamiltonian with momentum mapping $J = x^1 y_2 - x^2 y_1 = W$.
Substituting $U = -XY,\ V = 1 - X^2$ and $W = l$ in the first equation we obtain

$$X^2Y^2 + l^2 = (1 - X^2)(Z - Y^2)$$

and hence

$$(1 - X^2)Z = Y^2 + l^2$$

and $|X| < 1$ as $V \geqslant 0$.

Consequently the set of $S^1 \approxeq SO(2)$-orbits in $J^{-1}(l)$ are seen to be in 1-1 correspondance with the set

$$M_l = \{(X, Y, Z) | (1 - X^2)Z = Y^2 + l^2, -1 \geqslant X \geqslant 1\}$$

If $l \neq 0$, then $|X| < 1$ and M_l is as the graph of the function $Z = \frac{Y^2+l^2}{1-X^2}$ and as such it is diffeomorphic to $\mathbb{R}^2$.

As for $l = 0$, we note that although M_0 is still homeohophic to $\mathbb{R}^2$ it is not the graph of a fuction in $\mathbb{R}^3$ since it contains the vertical lines $(1, 0, Z)$ and $(-1, 0, Z)$.

Now let us compute the Poisson structure on T^*S^2/S^1. X, Y, Z being invariant under the action of $S^1 \approxeq SO(2)$, their Poisson brackets on T^*S^2 will be sufficient to describe the dynamics on M_l. Since the matrix $[\{F_i, F_j\}]_{1 \leq i,j \leq 2}$ is invertible the Poisson bracket on T^*S^2/S^1 may be computed using the Dirac process.([Dir]). It can be checked that

$$\begin{aligned} \{X, Y\}_{T^*S^2/S^1} &= 1 - X^2 \\ \{X, Z\}_{T^*S^2/S^1} &= 2Y \\ \{Y, Z\}_{T^*S^2/S^1} &= -2XZ \end{aligned} \tag{6.1}$$

It is interesting to note, that putting $X^1 = X,\ X^2 = Y,\ X^3 = Z$ and setting $\psi(X, Y, Z) = Z(1 - X^2) - Y^2 - l^2$ the equations (6.2)can be written in the form

$$\{X^i, Xj\} = \epsilon_{ijk} \frac{\partial \psi}{\partial X^k} \tag{6.2}$$

By using (6.2) we can deduce that

$$\{f, g\} = (\nabla f \times \nabla g) \cdot \nabla \psi$$

Therefore Hamiltonian equations for $H \in C^{\infty}(M_l)$ can be written as

$$\frac{dX^i}{dt} = (\nabla H \times \nabla \psi)_i$$

Now putting

$$\begin{aligned} H &= \frac{1}{2}\langle \mathsf{y}, \mathsf{y} \rangle + \gamma x^3 \\ &= \frac{1}{2} Z + \gamma X \end{aligned}$$

We can write the equations of motion as

$$\begin{aligned} \frac{dX}{dt} &= \{X, H\} = Y \\ \frac{dY}{dt} &= \{Y, H\} = -\gamma(1 - X^2) - XZ \\ \frac{dZ}{dt} &= \{Z, H\} = -2\gamma Y \end{aligned}$$

CHAPTER 7

MAGNETIC SPHERICAL PENDULUM

Consider the $S^1 = SO(3)$- action on $T^*\mathbb{R}^3$ defined in the previous chapter. Working with the symplectic form

$$\Omega = \omega + \frac{\mu}{\|\mathsf{x}\|^3}\epsilon_{ijk}x^i dx^j \otimes dx^k$$

it can be checked that this action is Hamiltonian and has a momentum mapping L

$$\begin{aligned} L(\mathsf{x},\mathsf{y}) &= \langle \mathsf{x}\times\mathsf{y},\mathsf{e}_3\rangle + \mu\langle\frac{\mathsf{x}}{|\mathsf{x}|},\mathsf{e}_3\rangle \\ &= x^1 y_2 - x^2 y_1 + \mu\frac{x^3}{\|\mathsf{x}\|^3} \end{aligned}$$

Indeed it can be checked that

$$\begin{aligned} (\frac{\partial}{\partial t})^* &= -x^2\frac{\partial}{\partial x^1} + x^1\frac{\partial}{\partial x^2} - y_2\frac{\partial}{\partial y_1} + y_1\frac{\partial}{\partial y_2} \\ &= X_J \end{aligned}$$

Again employing the invariant polynomials $X,\ Y,\ Z,\ U,\ V,\ W$ introduced in the previous chapter we remember that they satisfy the relations

$$\begin{aligned} U^2 + W^2 &= V(Z - Y^2) \\ Z &\geqslant 0 \\ V &\geqslant 0 \end{aligned}$$

Now working with the symplectic form Ω we find the non-zero Poisson brackets to be

$$\{X, Y\} = 1$$

$$
\begin{aligned}
\{X,Z\} &= 2Y \\
\{Y,Z\} &= \frac{2\mu}{|x|^3}W \\
\{Y,V\} &= \frac{\mu}{|x|^3}V \\
\{Z,U\} &= \frac{2\mu}{|x|^3}XW + 2Y^2 - 2Z \\
\{Z,V\} &= -4U \\
\{Z,W\} &= -\frac{2\mu}{|x|^3}(XU+YV) \\
\{U,V\} &= 2V \\
\{U,W\} &= -\frac{\mu}{|x|^3}XV
\end{aligned}
$$

For the points on $T^*S^2 \subseteq T^*\mathbb{R}^3$ we again have

$$
\begin{aligned}
V + X^2 &= 1 \\
U + XY &= 0
\end{aligned}
$$

and for the points on $T^*S^2 \cap L^{-1}(l)$ we have

$$(1 - X^2)Z = Y^2 + W^2$$

where $|X| < 1$ and $Z \geqslant 0$.

However, in this case

$$l = L = W + \mu X$$

and

$$(1 - X^2)Z = (l - \mu X)^2 + Y^2$$

with $|X| < 1$ and $Z \geqslant 0$.

When $j \neq \pm\mu$, the set of orbits of $S^1 \approxeq SO(2)$ in $T^*S^2 \cap L^{-1}(l)$ is diffeomorphic to $\mathbb{R}^2$. It is indeed the graph of function

$$Z = \frac{(j - \mu X)^2 + Y^2}{1 - X^2}, \qquad |X| < 1$$

When $j = \pm\mu$, M_j is still homeomorphic to $\mathbb{R}^2$, although M_j is not the graph of a function, because it contains vertical lines $\{(\pm 1, 0, Z) \in \mathbb{R}^3 | Z \geqslant 0\}$.

$H|_{T^*S^2}$ induces the Hamiltonian $H_j|_{M_j}$ where

$$\begin{aligned} H_j : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (X, Y, Z) &\longmapsto \frac{1}{2}Z + \gamma X \end{aligned}$$

Now we compute the Poisson structure on $C^\infty(M_j)$. Recall that a Poisson bracket $\{\cdot, \cdot\}_{T^*\mathbb{R}^3}$ on $C^\infty(T^*\mathbb{R}^3)$ is

$$\{f, g\}_{T^*\mathbb{R}^3} = \frac{\partial f}{\partial \xi_i} \frac{\partial g}{\partial \xi_j} \{\xi_i, \xi_j\}_{\mathbb{R}^3}$$

where $(\xi_i)_{1 \le i \le 6}$ are coordinates of $T^*\mathbb{R}^3$. Then the structure matrix $[\{\xi_i, \xi_j\}]_{1 \le i,j \le 6}$ is a skew- symmetric matrix whose nonzero elements are listed below:

$$\begin{aligned} \{x^1, y_1\} &= \{x^2, y_2\} = \{x^3, y_3\} = 1 \\ \{y_1, y_2\} &= -\mu \frac{x^3}{|x|^3} \\ \{y_1, y_3\} &= \mu \frac{x^1}{|x|^3} \\ \{y_2, y_3\} &= -\mu \frac{x^2}{|x|^3} \end{aligned}$$

$$\{F_1, F_2\} = \{V + X^2, U + XY\}_{\mathbb{R}^6}|_{T^*S^2/S^1} = 2$$

Therefore, the Poisson bracket on T^*S^2/S^1 may be computed using the Dirac prescription. Recall that we can write Hamiltonian equations for $H \in C^\infty(M_l)$ as

$$\frac{dX^i}{dt} = (\nabla H \times \nabla \psi)_i$$

In particular for the Hamiltonian function

$$H = \frac{1}{2}Z + \gamma X$$

we obtain the governing equations for the motion of the magnetic spherical pendulum as

$$\begin{aligned} \frac{dX}{dt} &= Y \\ \frac{dY}{dt} &= -\gamma(1 - X^2) + \mu(j - \mu X) - XZ \\ \frac{dZ}{dt} &= -2\gamma Y \end{aligned}$$

CHAPTER 8

CONCLUSION

Our work on the governing equations of quite standard objects of classical mechanics seems to indicate that much clarity can be achieved by expressing such problems within the framework of symplectic manifolds. In this respect the importance of the ideas of P. M. Dirac is quite striking. Finally the Poisson formalism represents the most efficient category for higher classical mechanics.

REFERENCES

[A-C-G] J. Arms, R. Cushman, M. Gotay : A universal reduction procedure for Hamiltonian group actions in *The Geometry of Hamiltonian Systems*, Springer- Verlag(**1991**)

[C-B 1] R. Cushman, L. Bates : The magnetic spherical pendulum. *Meccanica* 30 (**1995**)271-289

[C-B 2] R. Cushman, L. Bates : *Global Aspects of Classical Integrable Systems*, Birkhäuser Verlag (**1997**)

[Dir] P. A. M. Dirac : *Lectures on Quantum Mechanics*, Academic Press (**1964**)

[L-M-S] E. Lerman, R. Montgomery, R. Sjamaar : Examples of singular reduction, *Symplectic Geometry*, London M. S. Lecture Notes Series (**1993**)

[Pars] L. A. Pars : *A Treatise on Analytical Dynamics*, Wiley(**1965**)

[Weyl] H. Weyl : *The Classical Groups*, Princeton University Press(**1946**)

[Whi] E. T. Whittaker : *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Cambridge [Eng.] The University Press (**1937**)

APPENDIX A

Theorem A.1 *Let (M, Ω) be a symplectic manifold and N be a submanifold of M where $N = F^{-1}(q)$ for some regular value $q \in M$ of $F = (F_1, F_2, \dots, F_r) : M \to \mathbb{R}^r$. $\Omega|_N$ is a symplectic form on N if the matrix*

$$\Big[\ \{F_i, F_j\} \ \Big]_{1 \leq r, j \leq r}$$

is non- singular at each point of N.

Proof: Consider $p \in N$. Let V_pM be the subspace of T_pM spanned by the vectors $X_{F_1}|_p, X_{F_2}|_p, \dots, X_{F_r}|_p$. If $u \in T_pN \leq T_pM$, then

$$\omega(X_{F_k}, u) = uF_k = 0$$

Consequently $T_pN \leq V_pM^{\perp\omega}$. On the other hand

$$dim T_pN = dim T_pM - dim \cap_{k=1}^{r} ker \ dF_k = dim T_pM - r = dim V_pM^{\perp\omega}$$

as ω is non-degenerate. We conclude that

$$T_pN = V_pM = \langle X_{F_1}|_p, \dots, X_{F_r}|_p \rangle.$$

and $\omega|_{T_pN}$ is non-singular as

$$\omega(X_{F_i}, X_{F_j}) = \{F_i, F_j\}.$$

APPENDIX B

Theorem B.1 *For any* $f, g \in C^\infty(M)$

$$\{f|_N, g|_N\}_N = \{f, g\}^*|_N$$

where bracket $\{\cdot,\cdot\}^*$ *on* $C^\infty(M)$ *is defined by*

$$\{f, g\}^* = \{f, g\} - \{f, F_i\}\Psi^{ij}\{F_j, g\}$$

with the matrix

$$\left[\ \Psi^{ij}\ \right]_{1\le i,j\le r}$$

being the inverse of the matrix

$$\left[\ \{F_i, F_j\}\ \right]_{1\le i,j\le r}$$

Proof: First observe that for any $f \in C^\infty(M)$, if we define

$$f^\star = f - \{f, F_i\}\psi^{ij}F_j$$

we have

$$\{f^*, F_k\} = \{f, F_k\} - \{\{f, F_i\}\psi^{ij}F_j, F_k\}$$

and

$$\begin{aligned}\{f^*, F_k\}|_N &= \{f, F_k\}|_N - \{\{f, F_i\}, F_k\}|_N\psi^{ij}F_j - \{\{f, F_i\}, \{\psi^{ij}, F_k\}|_N F_j \\ &- \{f, F_i\}|_N\psi^{ij}\{F_j, F_k\}|_N \\ &= \{f, F_k\}|_N - \{f, F_i\}|_N\delta^i_k = 0\end{aligned}$$

Consequently N is invariant under X_{f^*} for any $f \in C^\infty(M)$. Clearly

$$X_{f^*}|_N = X_{f^*|_N}$$

Secondly we notice that for any $f, g \in C^\infty(M)$

$$\begin{aligned}
\{f^*, g^*\}|_N &= \{f - \{f, F_i\}\psi^{ij}F_j, g - \{\{g, F_k\}\psi^{kl}F_l\}|_N \\
&= \{f, g\}|_N - \{g, F_k\}\psi^{kl}\{f, F_l|_N - \{f, F_i\}\psi^{ij}\psi^{kl}\{F_j, F_l\}|_N \\
&+ \{f, F_i\}\{g, F_k\}\psi^{ij}\psi^{kl}\{F_j, F_l\}|_N \\
&= \{f, g\}|_N - \{f.F_i\}\psi^{ij}\{F_j, g\}|_N \\
&= \{f, g\}^*|_N
\end{aligned}$$