Nonholonomic Dynamical Systems

A thesis presented in partial
fulfilment of the requirements

for the degree

of Master of Science

in Mathematical Physics at
Massey University

Brett Ryland
2002

Abstract

The dynamics of mechanical systems subject to nonholonomic (i.e. non-integrable velocity) constraints is only poorly understood. It is known that (i) they preserve energy and, (ii) they are reversible. In this thesis I explore the conjecture that (i) and (ii) are the *only* general features of the entire class. The discovery of dissipative orbits, ones that behave differently as $t \to +\infty$ and $t \to -\infty$, would strongly support this conjecture.

This dissipation can appear in various forms, e.g. sinks (attractors) or sources (repellers) in the phase space, but in every form the dynamics have the property that the forwards time orbit occupies a different region of the phase space than the reverse time orbit.

In nonholonomic dynamical systems that are reversible and possess an integral, theory predicts that near the fixed set of a reversing symmetry, e.g. $R : p \mapsto -p$ with fixed set $\mathrm{Fix}(R) = \{(q,p) : p = 0\}$, no dissipation can occur. If the system can be integrated analytically, then all the orbits are quasi-periodic and even away from the fixed set of any reversing symmetries, dissipation cannot occur. But, if the system cannot be integrated analytically, then away from the fixed set of any reversing symmetries, dissipative orbits can exist.

The minimum dimension needed for a nonholonomic system is 6. So, in this thesis I study the simplest class of nonholonomic dynamical systems that are reversible with an integral, namely the contact particle in $\mathbb{R}^3$. I search for evidence of dissipative behaviour in this class of systems by taking a known contact particle system that can be integrated analytically, such as the harmonic oscillator, where no dissipation can occur and calculating (numerically and analytically) the dynamics of its orbits. Then I perturb the system so that it cannot be integrated analytically and search for orbits that exhibit the dissipative behaviour described above away from the fixed set of the reversing symmetries of the system.

To achieve this I implemented a semi-explicit reversible integrator in C to integrate the system forwards (or backwards when desired) in time from an initial point. The C code interacts with MATLAB via the "mex" interface to make use of MATLAB's graphing facilities, which I used to plot the forwards and backwards orbits in blue and red respectively. This allows the orbits to be observed and any dissipative behaviour should become immediately apparent as the orbits will cover different portions of the phase space if dissipation occurs. The phase space of the system is actually $\mathbb{R}^6$, which is beyond my capabilities to visualise, but it can be reduced to $\mathbb{R}^3$, as I have done, through the use of the integral, the nonholonomic constraint and a Poincaré section.

Acknowledgements

I would like to acknowledge my supervisor, Assoc. Prof. Robert McLachlan, and his postdoc, Matt Perlmutter, for their assistance with the theoretical aspects of nonholonomic dynamical systems. Also worthy of acknowledgement are the members of the Institute of Fundamental Sciences at Massey University for their knowledge of mathematics and physics in general and for their moral support. In addition, thanks to the graduate students in the Mathematics Department at Massey University whose antics are a constant source of entertainment.

Contents

List of Figures

Chapter 1

Introduction

1.1 Overview

In this thesis I am studying a class of simple mechanical systems subject to a single nonholonomic constraint.

A simple mechanical system is a dynamical system that has a Hamiltonian

$$H(q,p) = \frac{1}{2}p^T\mu(q)^{-1}p + V(q)$$

where $\mu(q)$ is a mass matrix and $V(q)$ is a potential. The equations of motion are given by

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

A nonholonomic dynamical system is a dynamical system which has a non-integrable constraint, $A(q)p = 0$, on the velocities. A simple mechanical system subject to a single nonholonomic constraint, with $\mu(q) = Id$, can be written as

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -\nabla V(q) + A^T(q)\lambda \\ A(q)p &= 0. \end{aligned} \tag{1.1}$$

One property of simple mechanical systems is that the energy, $H(q,p)$, is preserved. This means that for a given initial point, $(q(0),p(0))$, in the system, with energy $H(q(0),p(0)) = h_0$, any point in the phase space that can be reached by integrating forwards, or backwards, in time from that initial point, must have the same energy, $H(q,p) = h_0$.

Another feature of simple mechanical systems is that they are reversible with respect to the reversing symmetry $R : (q,p) \mapsto (q,-p)$. So, if we apply R to the unconstrained system above, we get

$$\begin{aligned} \dot{q} &= -p \\ -\dot{p} &= -\nabla V(q) + A^T(q)\lambda, \end{aligned}$$

then, applying $t \mapsto -t$ returns us to the original system.

The dynamics of mechanical systems subject to nonholonomic constraints is poorly understood. In particular, it is not known whether or not these nonholonomic systems are Hamiltonian, or volume preserving. In this thesis I explore the conjecture that in the class of simple mechanical systems subject to nonholonomic constraints, the *only* general features of the dynamics is that they preserve the energy and are reversible.

The theory of nonholonomic dynamical systems predicts that for a system as described above, near the fixed set of a reversing symmetry the path of almost any point in the phase space is such that it will return to a neighbourhood of the initial point infinitely many times [32]. If the system can be integrated analytically, then all the orbits are quasi-periodic and even away from the fixed set of any reversing symmetries, the above holds. But, if the system cannot be integrated analytically, then away from the fixed set of any symmetries, this condition no longer applies and dissipation can occur. That is, the path of a point in phase space as we integrate forwards in time can occupy one region of the phase space while the path of the same point can occupy a different region of the phase space when we integrate backwards in time.

The discovery of dissipative behaviour in a nonholonomic dynamical system would strongly support the above conjecture. It can be shown that the lowest dimension needed for a nonholonomic dynamical system is 6. Thus the simplest class of nonholonomic dynamical systems is the "contact particle" in $\mathbb{R}^3$, defined by the Hamiltonian, $H(q,p) = \frac{1}{2}\|p\|^2 + V(q)$, and the constraint $p_1 + q_2 p_3 = 0$. It is in this class of systems that I am searching for evidence of dissipative behaviour.

1.2 History and Background

The word 'holonomic' is due to Hertz and means 'universal' or 'integrable', literally it translates as '$\acute{o}\lambda o\varsigma$' = 'whole', '$\nu o\mu o\varsigma$' = 'law' [36]. Thus the term 'nonholonomic' is synonymous with 'non-integrable'. The theory of nonholonomic systems is the subject of many papers, some of which date back to the turning of the last century, but the term 'nonholonomic' is scarcely even mentioned in most texts.

Nonholonomic variational problems have much in common with optimal control problems and occur in thermodynamics and quantum theory. They are also closely connected with the general theory of partial differential equations.

The beginnings of nonholonomic theory can be traced back to Lagrange in 1788, with his *Méchanique Analytique* [22], in which the equations of unconstrained motion are written in Euler-Lagrange form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\right) = \frac{\partial L}{\partial q^i},$$

in section V. He used the notation $Z = T - V$ for his Lagrangian and concluded that a coordinate invariant expression for *mass* $\times$ *acceleration* is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v}\right) - \frac{\partial T}{\partial q}.$$

Curiously [26], Lagrange does not recognise the equations of motion as being equivalent to the variational principle

$$\delta \int L dt = 0,$$

despite knowing the general form of the differential equations for variational problems; he had actually commented on Euler's proof of them. This was only recognised a few decades later by Hamilton.

Although Euler had already treated nonholonomic constraints [17], they were not clearly understood until the turn of last century. Foremost in clarifying the features of nonholonomic mechanical systems was Hertz with his *Prinzipien der Mechanik* in 1894. Gibbs and Carathéodory also dealt with contact structures in the formulation of thermodynamics. As for pure mathematics, the study of nonholonomic systems began with the theory of Pfaffian systems and the subsequent work on the general theory of differential equations. E. Cartan introduced the powerful tools of differential forms and codistributions but, unfortunately, these were not widely used in nonholonomic problems.

In the 1920's Levi-Civita and H. Weyl defined the notions of Riemannian and affine connections and discovered deep relations between mechanics and geometry. Nonholonomic mechanics served as a source of new geometrical structures which provided mechanics and physics with a convenient and concise language. This mutual interaction was started in the pre-war years by Vranceanu and Synge. In 1931, Vranceanu [37] gave the first precise definition of a nonholonomic structure on a Riemannian manifold in two short notes and an article, outlining its relation to dynamics of nonholonomic systems. Synge [34, 35] studied the stability of free motion of nonholonomic systems and anticipated the notion of curvature of a manifold.

According to Vershik and Gershkovich [36], in the post-war years the research on nonholonomic systems waned due to the vagueness of how the papers were written, giving vast differences in the notations and coordinates used, resulting in large, mostly incomprehensible texts. This was emphasised in 1948 when V. Vagner wrote: "The lack of rigour which is typical for differential geometry is reflected also in the absence of precise definitions of such notions as spaces, multi-dimensional surfaces, etc. Differential geometry is certainly dropping behind and this became even more dangerous when it lost its direct contact with theoretical physics" [36]. In the 1950's and 1960's nonholonomic theory was almost completely left alone.

In 1975, Vershik and Faddeev produced a paper in which nonholonomic mechanics was exposed systematically in terms of differential geometry. With the introduction of a more consistent notation, there came a renewed interest in nonholonomic theory.

More recently, the amount of interest in nonholonomic theory has increased significantly, with several papers in the late 1990's on symmetries, reduction and the application of nonholonomic theory to molecular dynamics and control theory, as I now briefly survey.

In a paper in 1995, Sarlet, Cantrijn and Saunders [33] discuss the concepts of symmetries and adjoint symmetries for Lagrangian systems with nonholonomic constraints. The following year Bloch, Krishnaprasad, Marsden and Murray [7] developed the geometry and dynamics of mechanical systems with nonholonomic constraints and symmetries from the perspective

of Lagrangian mechanics with a view to control-theoretical applications. In particular they derived the evolution equation for momentum and distinguished geometrically and mechanically between the cases where it is conserved and where it is not. Also in 1996, Cardin and Favretti [14] derived the vakonomic equations (equations of motion of nonholonomic systems, originally derived by Arnold, Kozlov and Neishtadt [2] in 1988) from a nonholonomic variational problem and gave a geometrical interpretation of the terms.

Koon and Marsden [20] wrote a paper, in 1998, comparing the Hamiltonian and Lagrangian formulations of nonholonomic systems. They further developed the theory of Poisson reduction, which is important in stability theory for nonholonomic systems and tied it to other work in the area. In another paper in the same journal, Cantrijn, de Leon, Marrero and de Diego [13] presented a geometric reduction procedure for Lagrangian systems. In 1999, Cortes and de Leon [15] developed a reduction scheme in terms of the nonholonomic momentum mapping.

Earlier in 1999, Cantrijn, de Leon and de Diego [12] unified several approaches to the 'almost-Poisson' bracket for mechanical systems with nonholonomic constraints and used the almost-Poisson structure to describe phase-space dynamics of a nonholonomic system. In a paper on molecular dynamics simulations, Kutteh [21] described three algorithms for imposing nonholonomic constraints with any number of additional holonomic constraints.

The theory of nonholonomic systems continues to develop, with strong links to other fields, such as molecular dynamics, thermodynamics, control theory and quantum mechanics, and is an increasingly popular field of research.

1.3 Definitions

Throughout this thesis I will be discussing simple mechanical dynamical systems subject to nonholonomic constraints, so a few definitions are in order.

Definition 1 (Dynamical System) *A* continuous time dynamical system *is a system of ODEs*

$$\dot{x} = F(x)$$

where $F(x)$ is a vector field on a manifold M, x is a point of the manifold and $\dot{x}$ denotes the time derivative of x.

For a finite dimensional manifold the system is representable by a set of first order ODEs, giving us a finite dimensional dynamical system. The solutions to the dynamical system are integral curves of F, which constitute a flow on the underlying manifold.

Definition 2 (Mapping) *A* mapping *is a diffeomorphism $f : M \mapsto M$ and is the discrete time analog of the flow of a vector field. Typically mappings are written as*

$$x_{i+1} = f(x_i) \qquad or \qquad x' = f(x).$$

Many ODEs can be reduced to a mapping via the *Poincaré* (or *first return*) *map*, which can be defined as follows:

Definition 3 (Poincaré Map) *Consider a dynamical system, $\dot{x} = F(x)$, in a manifold, M. Then a* Poincaré section, Σ, *can be defined as follows:*

Let $\Sigma \subset M$ be of co-dimension 1 such that F is transverse to Σ, i.e. $F(x) \notin T_x\Sigma$. Then Σ is a Poincaré section if, for all $x \in \Sigma$, there exists a minimum $t = \tau > 0$ such that $\phi_\tau(x) \in \Sigma$ and $\phi_\tau(x)$ crosses Σ in the correct sense.

From the Poincaré section we can define the Poincaré map, $\phi : \Sigma \mapsto \Sigma$, *as*

$$\phi(x) := \phi_\tau(x).$$

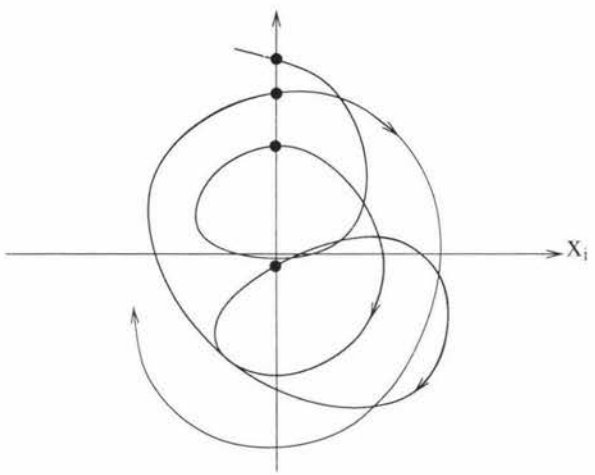


Figure 1.1: An example of a Poincaré map. The dots indicate where the trajectory crosses $x_i = 0$ in the positive direction.

For example, on $M = \mathbb{R}^n$,

$$\Sigma = \{x : x_i = 0, \dot{x}_i > 0\}$$

is a Poincaré section if all points in Σ eventually return to Σ.

Definition 4 (Orbit) *For the system $\dot{x} = F(x)$ with flow $x(t) = \phi_t(x(0))$, an* orbit *(or* trajectory*) $\mathcal{O}(x)$ passing through a point x is the set of points in the phase space reachable from x by flowing (or mapping in the case of maps) forwards or backwards in time.*

i.e.

$$\mathcal{O}(x) = \{\phi_t(x) : t \in \mathbb{R}\} \qquad or \qquad \mathcal{O}(x) = \{f_i(x) : i \in \mathbb{Z}\}.$$

If an orbit passing through a point x^* returns to that point after a time t^*, then the orbit is called *periodic* of period t^*. If the orbit can be written as

$$\mathcal{O}(x) = g(\omega_1 t, \ldots, \omega_k t) \text{ with } \frac{\omega_i}{\omega_j} \text{ irrational for } i \neq j$$

and g being 2π-periodic in each argument, then $\mathcal{O}(x)$ is said to be *quasi-periodic* and the phase space covered by the orbit, $g : \mathbb{T}^k \mapsto M$, is a *k-torus.*

Definition 5 (Fixed Point) *A* fixed point x^* *of a system is an orbit which remains at* x^* *for all time* t.

In the case of maps, if a point x^* in an orbit returns to x^* after q steps, where $q \geq 1$ is the smallest such integer, then the point x^* is called a *periodic point* of *period* q and the orbit $\mathcal{O}(x) = \{x^*, f(x^*), \ldots, f_{q-1}(x^*)\}$ is called a *q-cycle.* So a fixed point in a map is really a periodic point of period 1.

Definition 6 (Fixed Set) *Let* R *be a map* $R : M \mapsto M$. *Then the* fixed set *of* R *is*

$$\text{Fix}(R) = \{x : R(x) = x\}.$$

That is, the set of points in the phase space that remain where they are when R *is applied.*

Definition 7 (Hamiltonian) *A* Hamiltonian *system is one in which there exists a twice continuously differentiable function* $H(q,p)$, *such that the system can be written as*

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad i = 1, \ldots, m$$

Definition 8 (non-Hamiltonian) *Any system that is not Hamiltonian.*

Definition 9 (Simple Mechanical) *A* simple mechanical *system is a Hamiltonian system in which the Hamiltonian is of the form*

$$H(q,p) = \frac{1}{2} p^T \mu(q)^{-1} p + V(q)$$

where $\mu(q)$ *is a symmetric mass matrix and* $V(q)$ *is a potential.*

Definition 10 (Volume Preserving) *A system is said to be* volume preserving *if there exists a volume element* $m(x)\mathrm{d}^n x$ *such that*

$$\det(\mathrm{D}\phi(x)) = \frac{m(x)}{m(x')}$$

for maps, or

$$\nabla \cdot (mF) = 0$$

for flows.

For a fixed point, $x' = x$, of a map the eigenvalues of $\mathrm{D}\phi(x)$ have product 1.

Definition 11 (Reversible) *A* reversible *system is a dynamical system with a reversing symmetry $R : M \mapsto M$ that reverses the direction of time. That is, if*

$$\frac{\mathrm{d}x}{\mathrm{d}t} = F(x) \qquad \textit{for all } x \in M$$

then

$$\frac{\mathrm{d}(R(x))}{\mathrm{d}t} = -F(R(x)),$$

i.e.

$$\mathrm{T}R.F(x) = -F(R(x)).$$

For example, consider the Hamiltonian

$$H(q,p) = \frac{1}{2}\|p\|^2 + V(q)$$

with the reversing symmetry

$$R : (q,p) \mapsto (q,-p).$$

The Hamiltonian describes the system

$$\begin{aligned} \dot{q} &= p, \\ \dot{p} &= -\nabla V(q). \end{aligned}$$

Applying R we get $p \mapsto -p$ giving the system

$$\begin{aligned} \dot{q} &= -p, \\ \dot{p} &= \nabla V(q). \end{aligned}$$

Then letting $t \mapsto -t$ we regain the original system, as can be seen in Figure 1.2. The reversing symmetry together with the reversal of time leaves the equations invariant.

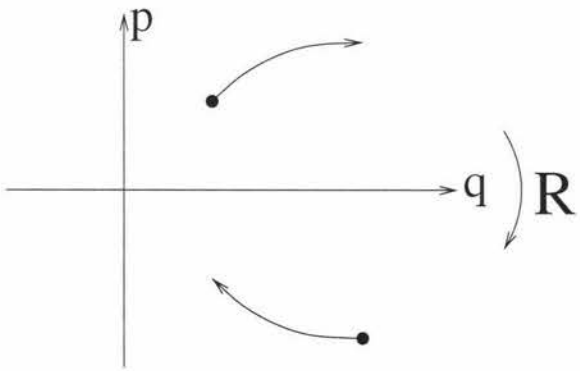


Figure 1.2: A point (q,p) being mapped forwards in time is equivalent to R being applied, that point being mapped backwards in time and then R being applied again.

Definition 12 (Nonholonomic) *[27] A* nonholonomic *system is a dynamical system with non-integrable constraints on the velocities. That is, for an n-dimensional nonholonomic system with $n-k$ constraints, the constraints*

$$\sum_{j=1}^{n} A_{ij}(q)p_j = 0, \qquad i = 1, \ldots, n-k$$

cannot be put into the form

$$\sum_{j=1}^{n} \frac{\partial f_i}{\partial q_j} p_j = 0$$

for some functions $f_1, \ldots, f_{n-k}$.

Examples of nonholonomic systems are the rolling penny, and the rattleback (see section 3.4), which both have a no-slip rolling constraint.

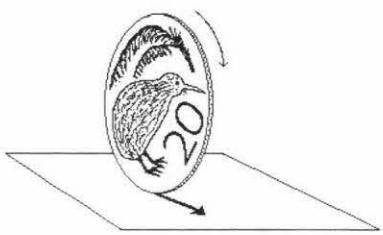


Figure 1.3: Rolling penny: the coin can only move in the direction it is rolling.

Figure 1.4 shows how dynamical systems can be classified into reversible mechanical systems, unconstrained mechanical systems and constrained mechanical systems. The constrained, reversible mechanical systems with nonholonomic constraints that I am discussing in this thesis are indicated.

1.4 Outline

As mentioned before, the discovery of dissipative behaviour in a non-integrable nonholonomic dynamical system would strongly support the conjecture that the only general features of the dynamics of the class of mechanical systems subject to nonholonomic constraints are that they preserve the energy and are reversible.

To find evidence of dissipative behaviour I am beginning with an unperturbed, simple Hamiltonian, which gives rise to an integrable, simple mechanical system and can be solved analytically. Then I compute the orbits of the system and find the period 2 points, as the amount of dissipation is easier to measure at fixed points, although no dissipation can occur yet. The code that I have written is a semi-explicit reversible integrator in C that interfaces with MATLAB via the "mex" interface to use MATLAB's graphing facilities. My code also

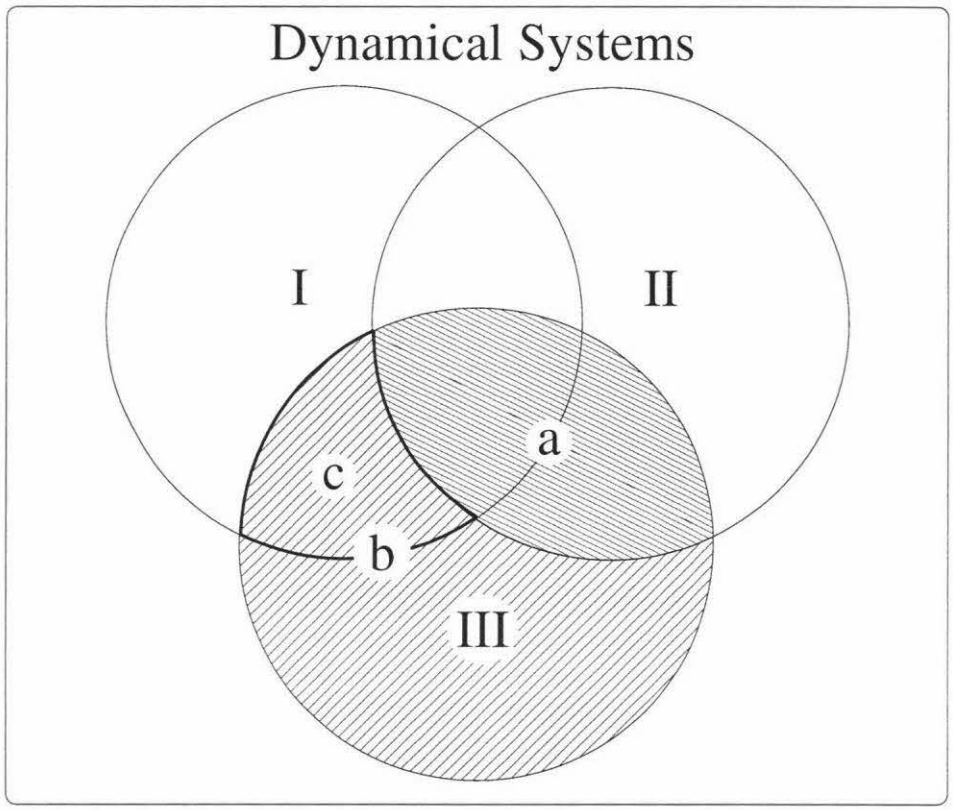


Figure 1.4: Dynamical systems.

I Reversible systems, $\frac{\mathrm{d}(R(x))}{\mathrm{d}t} = -F(R(x))$.

II Unconstrained mechanical systems. These are the systems defined purely by Euler-Lagrange / Hamiltonian equations of motion.

III Constrained mechanical systems.

a) The constraints are integrable, the system is said to be holonomic. It can be shown that the dynamics is equivalent to an unconstrained system on a lower dimensional manifold.

b) The constraints are non-integrable, the system is said to be nonholonomic.

c) The constrained, reversible mechanical systems with nonholonomic constraints that I am discussing in this thesis are contained in this region.

reduces the six dimensional system via the Hamiltonian integral, the nonholonomic constraint and a Poincaré section to three dimensions allowing the orbits to be plotted and visualised in MATLAB. The forwards time orbits are plotted in different colours to the reverse time orbits to make it immediately obvious when an orbit differs in forwards and backwards time.

Next, I perturb the system to make it non-integrable allowing for dissipation and I search for orbits that differ in forwards and backwards time. I also look for the period 2 points that survive the perturbation and measure the eigenvalues of the linearised coefficient matrix, which gives the amount of expansion in each direction at the periodic point and describes the amount of dissipation. Fixed points are not the only places where dissipation can occur, orbits such as invariant tori and some chaotic orbits can show dissipation but measuring it involves finding the eigenvalues of the Jacobian matrix of the map, which has to be calculated at each step and is more expensive in processor time.

The content of the remainder of this thesis is as follows. In chapter 2 I will discuss the history and properties of reversible dynamical systems, including Hamiltonian systems, maps and the linearisation and stability of fixed points.

Chapter 3 deals with the principle of nonholonomic mechanics, which is based on the Lagrange-d'Alembert principle and the associated Lagrange-d'Alembert equations. I will also cover some examples of nonholonomic dynamical systems, namely the simple, integrable, rolling penny, the complicated, high dimensional, rattleback and the contact particle with a spherically symmetric potential.

I will explain, in detail, how I went about searching for dissipative behaviour in the contact particle class of nonholonomic systems in chapter 4. This will involve the methods that I used to integrate the systems and my strategies for locating dissipative orbits, as well as the systems themselves and their perturbations. Also covered is the multitude of problems I encountered with the symmetries of the first system and the first Poincaré section and the resolution of these problems.

Finally, my results and conclusions will be discussed and summarised in chapter 5. The code that I have written is included in the appendices.

Chapter 2

Reversible Dynamical Systems

2.1 Some History of Reversible Dynamical Systems

A reversible dynamical system is defined in chapter 1 by definition 11. In short, a dynamical system is *reversible* if it has a reversing symmetry which, together with the reversal of time, leaves the equations invariant.

The following history is based on the accounts of Lamb and Roberts [24] and Roberts and Quispel [31].

Many of the differential equations of physics are time reversible. This was first noticed by Loschmidt [25], in 1877, for particles moving in a velocity independent force field. But, despite the physical laws being reversible, physical phenomena in general are not. This became known as Loschmidt's paradox when he pointed out that Boltzmann's second law of thermodynamics, which states that entropy is a monotonically increasing function of time, violates the time reversal symmetry of the microscopic equations of motion of the particles involved. The usual explanation is that although the laws of physics are invariant under time reversal, the initial conditions are such that time evolution is not.

Boltzmann [8, 9, 10] recognised the importance of time-reversing symmetry and showed that Maxwell's equations are reversible (if $B \mapsto -B$ as well). Also, Einstein's equations of classical general relativity [29, 30] are reversible and Birkhoff [5] utilised time reversal symmetry in 1915 during his studies of the restricted 3-body problem in classical mechanics. In particular, he noticed that if R was an involution ($R^2 = Id$) then the mapping f could be written $f = R \circ T$ where $R^2 = T^2 = Id$. After Birkhoff, time-reversal symmetry received little attention until the 1960's when, in particular, Hale [18] described time-reversal symmetry as "property E" and Wigner [38] showed the importance of time reversal symmetry in quantum mechanics.

In 1976, Devaney [16] noted that dynamical systems with involutory reversing symmetries other than the anti-symplectic one, $R : (q, p) \mapsto (q, -p)$, also exhibited many of the consequences of time-reversal symmetry. From this he defined reversible systems as those in which

the involutory nature of R was central and that R fixes a subspace one half the dimension of the phase space.

Later Arnol'd and Sevryuk [3] relaxed the latter condition and allowed any involutory reversing symmetry. Also they remarked that systems with reversing symmetries need not have involutory reversing symmetries, e.g. in quantum mechanical systems. Those systems are called weakly reversible systems and have many of the same properties as reversible systems, often the two are not distinguished between and the nature of the reversing symmetry is not mentioned.

Historically, reversibility is a symmetry property that arises mostly in Hamiltonian systems, due to the vast majority of reversible dynamical systems that arise in mechanical systems being Hamiltonian. In classical mechanics, reversibility is a powerful tool that is used to study periodic, homoclinic and heteroclinic orbits.

2.2 Symmetries

If an odd number of reversing symmetries are composed together then the resulting map is also a reversing symmetry, but if an even number of reversing symmetries are composed together then the resulting map is a symmetry S, with the property

$$\frac{\mathrm{d}S(x)}{\mathrm{d}t} = F(S(x)).$$

The group of symmetries together with reversing symmetries form the reversing symmetry group G [23] and the symmetries form a normal subgroup $H \trianglelefteq G$ [24].

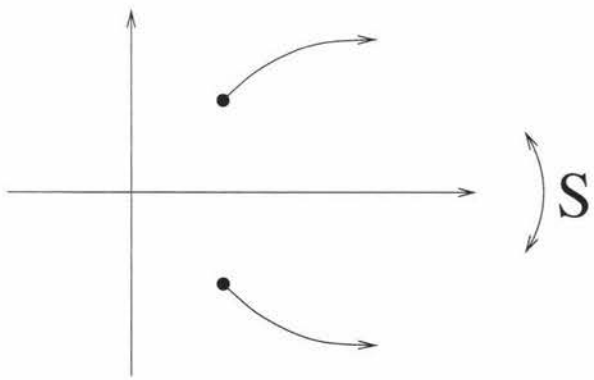


Figure 2.1: A point being mapped forwards in time is equivalent to S being applied, that point being mapped forwards in time and S being applied again.

2.3 Hamiltonian Systems

A *first integral* of $\dot{x} = F(x)$ on the region D is a continuously differentiable function $f : D(\subseteq \mathbb{R}^m) \mapsto \mathbb{R}$ in which $f(x(t))$ is a constant for any solution $x(t)$ of the system. Systems, such as Hamiltonian systems, which have a first integral on the whole space are *conservative*, typically the constant that is conserved is energy.

Some features of Hamiltonian systems are:

- H is a constant of the motion.
- The flow of a Hamiltonian system is symplectic [26].
- Hamiltonian systems preserve volume (or area in 2 dimensions).
- Poincaré recurrence: if the surfaces with constant H are compact then for almost all initial conditions, the system returns arbitrarily close to the initial point for arbitrarily long times [28], (discussed further in section 4.1.1 below).

As well as being classified as reversible and non-reversible, dynamical systems can be classified as Hamiltonian and non-Hamiltonian dynamical systems. As described in Definition 7, a system of differential equations on $\mathbb{R}^{2m}$ is a Hamiltonian system with m degrees of freedom if it can be written in the form:

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad i = 1 \ldots m$$

where $H(x) = H(q_1, \ldots, q_m, p_1, \ldots, p_m)$ is a twice continuously differentiable first integral.

The most common and important example of a reversible Hamiltonian dynamical system is one that has the property

$$H(q,p) = H(q,-p) \qquad \text{for all } q,p \in \mathbb{R}^n$$

which is reversible with respect to the conventional anti-symplectic reversing symmetry

$$R : (q,p) \mapsto (q,-p),$$

because then $\frac{\partial H}{\partial p}$ is odd in p and $\frac{\partial H}{\partial q}$ is even in p.

We emphasise that reversible systems need *not* be Hamiltonian. Non-Hamiltonian systems can show behaviour similar to Hamiltonian systems near symmetric orbits and dissipative or expansive behaviour near asymmetric orbits. Reversibility was taken more seriously as a symmetry property after many of the results for Hamiltonian systems were obtained by only assuming that the dynamical system was reversible, the most notable of which is KAM theory [24].

Some examples of non-Hamiltonian reversible systems are:

1.

$$\begin{aligned}\dot{x} &= zx + y + C_1\\ \dot{y} &= zy - x\\ \dot{z} &= C_2 z - x^2 - y^2\end{aligned}$$

where C_1 and C_2 are parameters. This simulates the dynamics of an externally injected class B laser where the damping rate associated with the field is much greater than that of the population, as in CO_2 lasers. The system is reversible with respect to the reversing symmetry $G : (x, y, z) \mapsto (-x, y, -z)$, but the divergence of the vector field is $2z$ and the dimension of the system is odd, hence the system is not Hamiltonian [31].

2. The evolution of a gas flame front in the presence of an external stabilising factor is described by a partial differential equation of the form

$$\xi_t + \xi_{xxxx} + 2\alpha\xi_{xx} + \xi + \beta\xi^3 + \gamma\xi^2 + \delta\xi_x^2 = 0$$

where $\xi = \xi(x, t)$ is the displacement of a flame point from its unperturbed position, x is a spatial coordinate, t is time and α, β, γ, and δ are parameters. The stationary solutions to the above equation can be written in the four-dimensional phase-space variables $(\xi, \xi_x, \xi_{xx}, \xi_{xxx})$ as a system of four coupled ordinary differential equations, which is invariant under $G : (\xi, \xi_x, \xi_{xx}, \xi_{xxx}) \mapsto (\xi, -\xi_x, \xi_{xx}, -\xi_{xxx})$ with $x \mapsto -x$. Thus it is reversible with x playing the role of time [31].

2.3.1 KAM Theory

KAM theory, named after its discoverers Kolmogorov, Arnol'd and Moser, deals with the persistence of invariant tori constituting quasi-periodic motion in nearly integrable dynamical systems.

It was originally developed in a Hamiltonian setting with smooth perturbations which preserve the symplectic structure but it was acknowledged that KAM theory applies to reversible dynamical systems early on.

Figure 2.2 shows some typical KAM circles surrounding an elliptic symmetric fixed point at the origin.

2.4 Mappings

For mappings, reversibility is defined by:

Definition 13 (Reversible Mapping) *The mapping $x \mapsto f(x)$ is* reversible *if there exists a mapping R such that $f \circ R = R \circ f^{-1}$.*

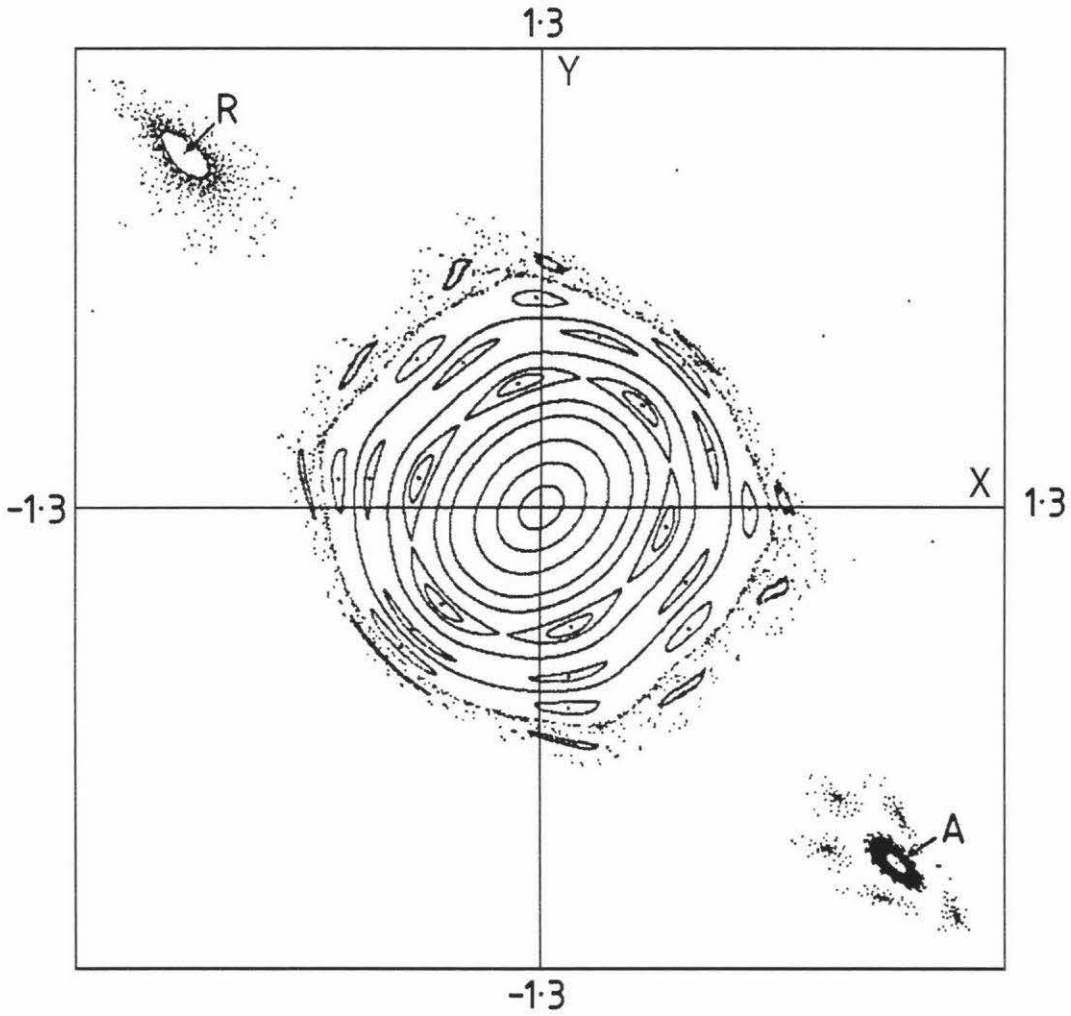


Figure 2.2: The phase portrait of a non-area-preserving reversible mapping of the plane with reversing symmetry $G : x' = y, y' = x$, showing conservative and dissipative behaviour. The origin is an elliptic symmetric fixed point and is surrounded by invariant curves. Elliptic symmetric 6-, 7-, and 8-cycles occur at the centres of concentric KAM circles, which are typically associated with conservative systems. Away from the origin, a trajectory spirals towards the attracting point A and its reflection spirals away from the repelling point, $\mathsf{R} = G\mathsf{A}$. An attracting 5-cycle is shown near A and many of the trajectories not enclosed by curves escape to infinity [31].

Thus reversible mappings have the property that $f \circ R \circ f = R$. Symmetries S are similarly defined by $f \circ S = S \circ f$. Also, if R is an involution then f can be written as the composition of two involutions.

There are some analogies between continuous-time systems and mappings. Hamiltonian ODEs are analogous to symplectic mappings, which have the property that they are volume-preserving (area-preserving in 2D) and KAM theories can be applied to them. Reversible DEs are analogous to reversible mappings and can arise from symmetric difference equations. In 2 dimensions, reversible mapping KAM theories imply that the neighbourhood of a symmetric periodic orbit is similar to the neighbourhood of a periodic orbit in an area-preserving mapping, under certain conditions. This can be seen in figure 2.2.

The determinant J of the Jacobian matrix of a mapping indicates whether the mapping is contracting $|J| < 1$, expanding $|J| > 1$, or area-preserving $|J| = 1$; the value of J can vary from point to point throughout the plane. Also, if $J > 0$ then the mapping is orientation-preserving (it preserves the order between sets of points on a curve), if $J < 0$ then the mapping is orientation-reversing (it reverses the order of points on a curve).

2.4.1 Examples

The most studied reversible mappings are the reversible symplectic mappings such as:

1. Henon mapping [31]

$$\begin{aligned} x' &= y \\ y' &= -x + 2Cy + 2y^2 \end{aligned}$$

which can be decomposed into the composition of two involutions $H \circ G$, with

$$H = \begin{cases} x' = x, \\ y' = -y + 2Cx + 2x^2, \end{cases} \qquad G : \begin{cases} x' = y, \\ y' = x. \end{cases}$$

2. Standard Mapping [31]

$$\begin{aligned} x' &= x + y \\ y' &= y + \tfrac{K}{2\pi} \sin(2\pi x') \end{aligned}$$

which can be decomposed into the composition of two involutions $H \circ G$, with

$$H = \begin{cases} x' = x, \\ y' = -y + \frac{K}{2\pi} \sin(2\pi x), \end{cases} \qquad G : \begin{cases} x' = x + y, \\ y' = -y. \end{cases}$$

The Poincaré (first return) map is another well studied mapping. It is a subtle question where the Poincaré map of a reversible system is itself reversible, which I will explore later in section 4.2.1.

2.5 R-Symmetric Orbits

An orbit $\mathcal{O}(x)$ which is set-wise invariant under a reversing symmetry R, i.e. $R \circ \mathcal{O}(x) = \mathcal{O}(x)$, is said to be *$R$-symmetric*; an orbit which does not remain invariant under R is said to be *asymmetric* (or *non-symmetric*) with respect to R.

2.5.1 Flows

Theorem 1 *If $\mathcal{O}(x)$ is an orbit of an autonomous vector field with time reversing symmetry R then:*

1 *If $\mathcal{O}(x)$ intersects* Fix(R) *then $\mathcal{O}(x)$ is symmetric with respect to R. In this case $\mathcal{O}(x)$ intersects* Fix(R) *in no more than 2 points.*

2 *$\mathcal{O}(x)$ intersects* Fix(R) *in precisely 2 points if and only if the orbit is periodic, not a fixed point and is symmetric with respect to R.*

Proof:

1 Suppose $\mathcal{O}(x)$ intersects Fix(R) at x_0. Then, since R is a reversing symmetry of the system, $\phi_t(x_0) = R \circ \phi_{-t}(x_0)$ and thus $\mathcal{O}(x)$ is symmetric with respect to R.

If $\phi_t(x)$ intersects Fix(R) at time $t = \tau_0$ and again at $t = \tau_0 + \tau$, then $R \circ \phi_t(x)$ must intersect Fix(R) in the same points at $t = \tau_0$ and $t = \tau_0 - \tau$ and thus cannot intersect Fix(R) at any other points, as shown in figure 2.3.

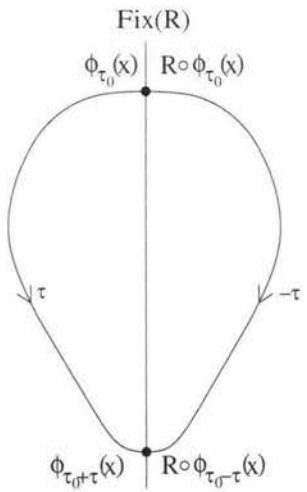


Figure 2.3: An R-symmetric orbit. The second point of intersection, found by flowing forwards by time τ, can also be found by flowing backwards by time τ and applying R.

2 Suppose $\mathcal{O}(x)$ intersects $\mathrm{Fix}(R)$ in precisely 2 points. Then $\mathcal{O}(x)$ is obviously not a fixed point and is symmetric with respect to R. If $\phi_t(x)$ intersects $\mathrm{Fix}(R)$ at time $t = \tau_0$ and again at $t = \tau_0 + \tau$, then as can be seen from figure 2.3 $\phi_t(x)$ must return to the initial intersection point at time $t = \tau_0 + 2\tau$ and is periodic.

Suppose $\mathcal{O}(x)$ is periodic, not a fixed point and is symmetric with respect to R, then, by part 1, $\mathcal{O}(x)$ intersects $\mathrm{Fix}(R)$ in no more the 2 points and since $\mathcal{O}(x)$ is not a fixed point and is periodic, $\mathcal{O}(x)$ intersects $\mathrm{Fix}(R)$ in more than 1 point. Therefore $\mathcal{O}(x)$ intersects $\mathrm{Fix}(R)$ in precisely 2 points.

2.5.2 Maps

Theorem 2 *[24] If $\mathcal{O}(x)$ is an orbit of an invertible map f with time reversing symmetry R then:*

1 *If $\mathcal{O}(x)$ intersects $\mathrm{Fix}(R) \cup \mathrm{Fix}(f \circ R)$ then $\mathcal{O}(x)$ is symmetric with respect to R. Then $\mathcal{O}(x)$ intersects $\mathrm{Fix}(R) \cup \mathrm{Fix}(f \circ R)$ in no more than 2 points.*

2 *$\mathcal{O}(x)$ intersects $\mathrm{Fix}(R) \cup \mathrm{Fix}(f \circ R)$ in precisely 2 points if and only if the orbit is periodic, not a fixed point and is symmetric with respect to R.*

- *$\mathcal{O}(x)$ has period $2p$ if and only if there exists $y \in \mathcal{O}(x)$ such that $y \in \mathrm{Fix}(R) \cap f^p\mathrm{Fix}(R)$ or $y \in \mathrm{Fix}(f \circ R) \cap f^p\mathrm{Fix}(f \circ R)$.*
- *$\mathcal{O}(x)$ has period $2p+1$ if and only if there exists $y \in \mathcal{O}(x)$ such that $y \in \mathrm{Fix}(R) \cap f^p\mathrm{Fix}(f \circ R)$.*

These properties are used in almost every paper on reversible dynamical systems as they justify searching for R-symmetric orbits in a subset of the full phase space. Also, they imply that in reversible dynamical systems, periodic orbits arise in families [24].

- In continuous m-parameter families of diffeomorphisms f_a with reversing symmetry R, R-symmetric even periodic orbits generically come in $(2\dim\mathrm{Fix}(R) - \dim\mathrm{Fix}(R^2) + m)$ and $(2\dim\mathrm{Fix}(f_a \circ R) - \dim\mathrm{Fix}(R^2) + m)$-parameter families.
- In continuous m-parameter families of autonomous flows with reversing symmetry R, R-symmetric periodic orbits of a given period typically arise in $(2\dim\mathrm{Fix}(R) - \dim\mathrm{Fix}(R^2) + m)$-parameter families.
- In continuous m-parameter families of autonomous flows with reversing symmetry R, R-symmetric periodic orbits typically arise in $(2\dim\mathrm{Fix}(R) - \dim\mathrm{Fix}(R^2) + m + 1)$-parameter families with smoothly varying period in which the fixed-period orbits are embedded.

2.6 Linearisation and Fixed Points

Definition 14 (Simple) *A fixed point is said to be* simple *if its linearised system is simple. A linearised system, $\dot{x} = Ax$, is* simple *if the matrix A is nonsingular, i.e. A has no zero eigenvalues.*

Near a fixed point the dynamics can often be described by a linearised system.

Theorem 3 (Hartman-Grobman Linearisation Theorem [4])
Let the nonlinear system

$$\dot{x} = F(x)$$

have a simple fixed point at $x = 0$. Then in a neighbourhood of the origin the phase portraits of the system and its linearisation are qualitatively equivalent provided the linearised system has no eigenvalues with zero real part.

So in the neighbourhood of the fixed point the nonlinear system looks like the linear system

$$\dot{x} = Ax$$

and the nature of the fixed point is determined by the eigenvalues of the linearised coefficient matrix A, i.e.

$$A_{ij} = \frac{\partial F_i}{\partial x_j}(0).$$

Reversible dynamical systems have an eigenvalue structure that is similar to that of Hamiltonian systems.

Theorem 4 *If λ is an eigenvalue of a symmetric fixed point of a reversible vector field then $-\lambda$ and $\bar{\lambda}$ are eigenvalues as well. If λ is an eigenvalue of a symmetric fixed point of a reversible diffeomorphism then λ^{-1} and $\bar{\lambda}$ are eigenvalues as well.*

Proof:

Suppose that in a system given by $\dot{x} = F(x)$ we have a fixed point x_0 of a reversing symmetry R. Then

$$\mathrm{d}R.F = -F(R). \tag{2.1}$$

Rewriting equation 2.1 in local coordinates we have

$$\frac{\partial R_i}{\partial x_j} F_j = -F_i(R(x))$$

which when differentiated gives

$$\frac{\partial R_i}{\partial x_j}\frac{\partial F_j}{\partial x_k} + \frac{\partial^2 R_i}{\partial x_j \partial x_k} F_j = -\frac{\partial F_i}{\partial x_j}\frac{\partial R_j}{\partial x_k}.$$

At the fixed point, $F(x_0) = 0$, leaving us with

$$\frac{\partial R_i}{\partial x_j}\frac{\partial F_j}{\partial x_k} = -\frac{\partial F_i}{\partial x_j}\frac{\partial R_j}{\partial x_k}.$$

or $R'F' = -F'R'$, which, upon rearranging, gives $R'F'R'^{-1} = -F'$. So $F'(x_0)$ is conjugate to $-F'(x_0)$ and must therefore have the same eigenvalues. The proof for maps is similar. Thus,

- For flows, the eigenvalues of symmetric fixed points come in singlets $\{0\}$, doublets $\{\lambda, -\lambda\}$ such that $\lambda \in \mathbb{R}$ or $\lambda \in i\mathbb{R}$, or quadruplets $\{\lambda, -\lambda, \bar{\lambda}, -\bar{\lambda}\}$.
- For reversible maps, the eigenvalues of symmetric fixed points come in singlets $\{\pm 1\}$, doublets $\{\lambda, \lambda^{-1}\}$ such that $\lambda \in \mathbb{R}$ or $\lambda \in S^1$, or quadruplets $\{\lambda, \lambda^{-1}, \bar{\lambda}, \bar{\lambda}^{-1}\}$.

2.6.1 Hyperbolic Fixed Points

Definition 15 (Hyperbolic Fixed Point) *A fixed point is* hyperbolic *if it has no eigenvalues with zero real part.*

If the fixed point of a linearised system is hyperbolic then, by the Hartman-Grobman linearisation theorem, the phase portrait of the nonlinear system is qualitatively equivalent to its linearisation in a neighbourhood of the fixed point. If we let n_u be the number of eigenvalues with $Re(\lambda) > 0$ and n_s be the number of eigenvalues with $Re(\lambda) < 0$ then $n_u + n_s = n =$ dimension of the system. Letting E^u and E^s be the unstable and stable eigenspaces of the matrix A, it is found that $\dim(E^u) = n_u$ and $\dim(E^s) = n_s$ where the elements of E^u have trajectories that approach the origin as $t \to -\infty$ and the elements of E^s have trajectories that approach the origin as $t \to \infty$.

Note that for nonlinear systems E^u and E^s have nonlinear analogs W^u and W^s, the unstable and stable manifolds and the elements approach the fixed point as $t \to -\infty$ and $t \to \infty$ respectively.

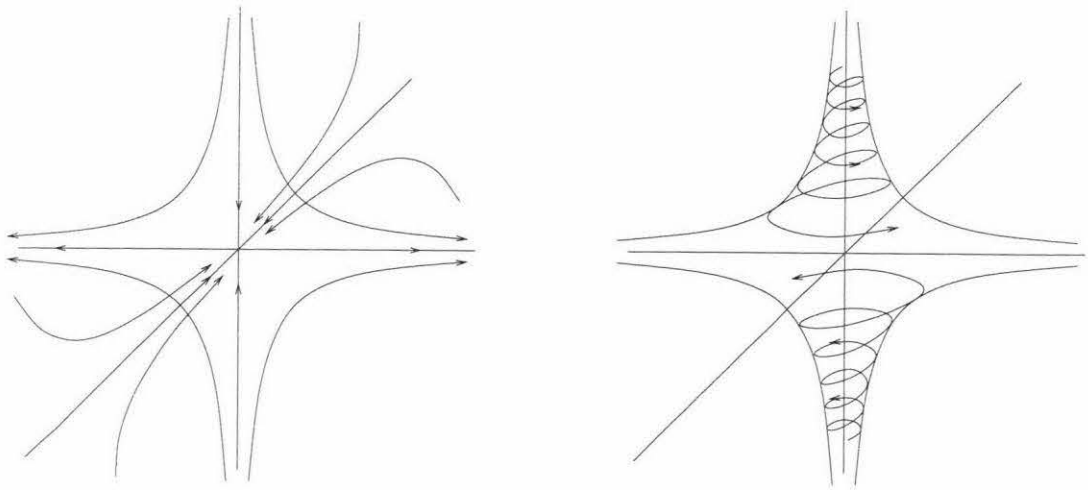
Some examples of hyperbolic fixed points are the node saddle point and the spiral saddle point in figure 2.4(a) and figure 2.4(b).

Non-hyperbolic fixed points can be non-simple and zero eigenvalues are allowed in the linearised coefficient matrix.

In the case of a Poincaré map, ϕ, a fixed point x^* corresponds to a closed orbit of the flow in the phase space. With coordinates $x_1, \ldots, x_m$, the linearisation, $\mathrm{D}\phi(x)$ of a Poincaré map $\phi : \mathbb{R}^m \mapsto \mathbb{R}^m$ is given by the $m \times m$ matrix

$$\mathrm{D}\phi(x) = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1} & \cdots & \frac{\partial \phi_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_m}{\partial x_1} & \cdots & \frac{\partial \phi_m}{\partial x_m} \end{bmatrix}$$

If no eigenvalue of $\mathrm{D}\phi(x^*)$ has modulus equal to unity then the fixed point of the Poincaré map and the closed orbit in the phase space are said to be *hyperbolic.*



(a) Node saddle point: the eigenvalues are $\lambda = \lambda_1, -\lambda_2, -\lambda_3$

(b) Spiral saddle point: the eigenvalues are $\lambda = \alpha \pm i\beta, -\lambda_3$

Figure 2.4: Two possible types of saddle points, the node saddle point and the spiral saddle point.

2.7 Stability

The stability of hyperbolic fixed points and periodic orbits are determined by linearised vector fields and return maps, hence R-symmetric periodic orbits cannot be asymptotically stable or unstable.

A fixed point x_0 of $\dot{x} = F(x)$ is considered stable if for every neighbourhood N of x_0 there exists another neighbourhood N' contained within N such that every trajectory passing through N' remains within N as time increases. There are three types of stability for fixed points. They are:

1. The point x_0 is *asymptotically stable* if it is stable and there exists a neighbourhood N of x_0 such that every trajectory passing through N approaches x_0 as $t \to \infty$.

2. The point x_0 is *neutrally stable* if it is stable but not asymptotically stable.

3. Any point which is not stable is *unstable*.

For the Poincaré map ϕ the stable (unstable) manifold $W^{s(u)}(x^*)$ of a fixed point x^* of ϕ is the set of points for which $\phi_k(x_0) \to x^*$ as $k \to \infty(-\infty)$. The stable (unstable) manifold of a periodic orbit is the union of the stable (unstable) manifolds of the fixed points of ϕ_q that constitute it.

2.8 Homoclinics and Heteroclinics

Homoclinics and heteroclinics typically form connections between saddle points. Chaotic behaviour is usually found in the neighbourhood of homoclinic and heteroclinic orbits. Letting $W^S_{x_0}$ and $W^U_{x_0}$ represent the stable and unstable manifolds of x_0, then:

- A point y is a *homoclinic* point of x_0 if y lies in the intersection of $W^S_{x_0}$ and $W^U_{x_0}$.
- A point y is a *heteroclinic* point of two fixed points x_1 and x_2 when $y \in W^S_{x_1} \cap W^U_{x_2}$.

These are difficult to find in general, but for reversible dynamical systems it is relatively easy to find symmetric homoclinic and heteroclinic orbits. If a dynamical system has a reversing symmetry R then:

- Let $y \neq x_0$ then $y \in \mathrm{Fix}(R) \cap W^{S,U}_{x_0}$ is a homoclinic point of x_0 if and only if x_0 is R-symmetric.
- Let $y \neq x_0$ then $y \in \mathrm{Fix}(R) \cap W^{S,U}_{x_0}$ is a heteroclinic point of x_0 if and only if x_0 is not R-symmetric.

Note that for maps, f, $\mathrm{Fix}(f \circ R)$ must also be considered.

Chapter 3

Nonholonomic Mechanics

3.1 Lagrange - d'Alembert Principle

The fundamental principle of nonholonomic mechanics is the Lagrange d'Alembert principle and the associated Lagrange d'Alembert equations.

A nonholonomic system can be defined by a Lagrangian $L(q, \dot{q})$ and a constraint. The Lagrange d'Alembert principle [26] states that if we consider the action functional given by

$$S(q(t)) = \int_0^t L(q(t), \dot{q}(t))\mathrm{d}t$$

over an interval of time I with fixed initial and final points $q_0, q_1 \in Q$ and look for critical points $q(t)$ of the action functional with respect to variations lying in the constraint distribution, which is given by

$$C_q = \{\dot{q} : \sum_{i=1}^{n} A_{ji}(q)\dot{q}_i = 0\}$$

then we will obtain a family of curves from which one curve, the orbit from q_0 to q_1, will satisfy

$$\dot{q}(t) \in C_{q(t)} \qquad \text{for all } t \in I.$$

In local coordinates, for an n-dimensional nonholonomic system with $n-k$ constraints we have the Lagrange d'Alembert equations [27]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = \sum_{j=1}^{n-k} \lambda_j A_{ji}(q) \qquad i \in \{1, \ldots, n\} \tag{3.1}$$

with the constraints:

$$\sum_{i=1}^{n} A_{ji}(q(t))\dot{q}_i = 0 \qquad j \in \{1, \ldots, n-k\}. \tag{3.2}$$

3.2 Discretised Lagrange - d'Alembert Principle

To allow numerical computation, the continuous equations of the Lagrange d'Alembert principle must be discretised. This can be done by a general-purpose DAE (differential algebraic equation) solver, or by discretising the Lagrange d'Alembert principle [27]. To do this we first choose a discrete Lagrangian $L_d : Q \times Q \mapsto \mathbb{R}$, which, letting $q_m \in Q$ be the position at time step m, gives us the discrete action sum:

$$S_d([q]) = \sum_{i=0}^{N} L_d(q_m, q_{m+1})$$

where $[q]$ is the discrete curve in Q. We then look for critical points of the discrete action sum with respect to variations of the discrete curve that lie in the discrete constraint distribution $C_d \subset Q \times Q$, which is a sub-manifold of dimension $n + k$.

Requiring $[q] \in C_d$ gives us the unique curve and we get the discrete Lagrange d'Alembert equations:

For each $i \in \{1, \ldots, N-1\}$

$$\frac{\partial S_d}{\partial q_{m,i}} = \sum_{j=1}^{n-k} \lambda_j A_{ji}(q_m)$$

with constraint $(q_m, q_{m+1}) \in C_d$. In chapter 4 I will use this to get a very fast, simple, reversible integrator for nonholonomic systems.

3.3 Properties of Nonholonomic Dynamics

In the field of nonholonomic dynamical systems almost nothing is known in general.

What is known is the following.

Theorem 5 *A simple mechanical system, which has a Lagrangian*

$$L = \frac{1}{2}\dot{q}^T \mu(q)\dot{q} - V(q), \tag{3.3}$$

is energy preserving for

$$H = \frac{1}{2}p^T \mu(q)^{-1} p + V(q)$$

where

$$p := \frac{\partial L}{\partial \dot{q}} = \mu(q)\dot{q},$$

and reversible under

$$R : (q, p) \mapsto (q, -p).$$

Proof:

Let $L = \frac{1}{2}\dot{q}^T\mu(q)\dot{q} - V(q)$ and $H = \frac{1}{2}p^T\mu(q)^{-1}p + V(q) = \frac{1}{2}\dot{q}^T\mu(q)\dot{q} + V(q)$. The Lagrange d'Alembert equations are

$$\begin{array}{rrcl} & \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} & = & \frac{\partial L}{\partial q} + A(q)^T\lambda \\ \Rightarrow & \frac{\mathrm{d}}{\mathrm{d}t}(\mu(q)\dot{q}) & = & \frac{1}{2}(\dot{q}^T\frac{\partial \mu}{\partial q})\dot{q} - \nabla V(q) + A(q)^T\lambda \\ \Rightarrow & (\dot{q}^T\frac{\partial \mu}{\partial q})\dot{q} + \mu(q)\ddot{q} & = & \frac{1}{2}(\dot{q}^T\frac{\partial \mu}{\partial q})\dot{q} - \nabla V(q) + A(q)^T\lambda \\ \Rightarrow & \mu(q)\ddot{q} & = & -\frac{1}{2}(\dot{q}^T\frac{\partial \mu}{\partial q})\dot{q} - \nabla V(q) + A(q)^T\lambda \end{array}$$

Now,

$$\begin{array}{rcl} \dot{H} & = & H_{\dot{q}}\ddot{q} + H_q\dot{q} \\ & = & (\mu(q)\dot{q})^T\ddot{q} + \frac{1}{2}\dot{q}^T(\frac{\partial \mu}{\partial q}\dot{q})\dot{q} + \nabla V(q)^T\dot{q} \\ & = & \dot{q}^T(\mu(q)\ddot{q}) + \frac{1}{2}(\dot{q}^T\dot{q})(\frac{\partial \mu}{\partial q}\dot{q}) + \nabla V(q)^T\dot{q} \\ & = & \dot{q}^T(-\frac{1}{2}(\dot{q}^T\frac{\partial \mu}{\partial q})\dot{q} - \nabla V(q) + A(q)^T\lambda) + \frac{1}{2}(\dot{q}^T\dot{q})(\frac{\partial \mu}{\partial q}\dot{q}) + \nabla V(q)^T\dot{q} \\ & = & -\frac{1}{2}(\dot{q}^T\dot{q})(\dot{q}^T\frac{\partial \mu}{\partial q}) + \frac{1}{2}(\dot{q}^T\dot{q})(\dot{q}^T\frac{\partial \mu}{\partial q}) - \dot{q}^T\nabla V(q)^T\dot{q} + \dot{q}^TA(q)^T\lambda \\ & = & \lambda^TA(q)\dot{q} \\ & = & 0 \end{array}$$

Thus, the system is energy preserving for $H = \frac{1}{2}p^T\mu(q)^{-1}p + V(q)$.

In terms of q and p, the Lagrangian can be written as

$$L = \frac{1}{2}p^T\mu^{-1}p - V(q),$$

and the Lagrange d'Alembert equations are given by

$$\begin{array}{rrcl} & \frac{\mathrm{d}}{\mathrm{d}t}(\mu(q)\dot{q}) & = & \frac{1}{2}(\dot{q}^T\frac{\partial \mu}{\partial q})\dot{q} - \nabla V(q) + A(q)^T\lambda \\ \Rightarrow & \frac{\mathrm{d}}{\mathrm{d}t}(p) & = & \frac{1}{2}((\mu^{-1}p)^T\frac{\partial \mu}{\partial q})(\mu^{-1}p) - \nabla V(q) + A(q)^T\lambda. \end{array}$$

Now, equation (11) from the definition of reversibility states

$$\frac{\mathrm{d}(R(x))}{\mathrm{d}t} = -F(R(x)),$$

and

$$\begin{array}{rcl} \frac{\mathrm{d}}{\mathrm{d}t}(-p) & = & -\frac{\mathrm{d}}{\mathrm{d}t}(p) \\ & = & -(\frac{1}{2}((\mu^{-1}p)^T\frac{\partial \mu}{\partial q})(\mu^{-1}p) - \nabla V(q) + A(q)^T\lambda) \\ & = & -(\frac{1}{2}((-\mu^{-1}p)^T\frac{\partial \mu}{\partial q})(-\mu^{-1}p) - \nabla V(q) + A(q)^T\lambda). \end{array}$$

So the system is reversible under $R : (q, p) \mapsto (q, -p)$.

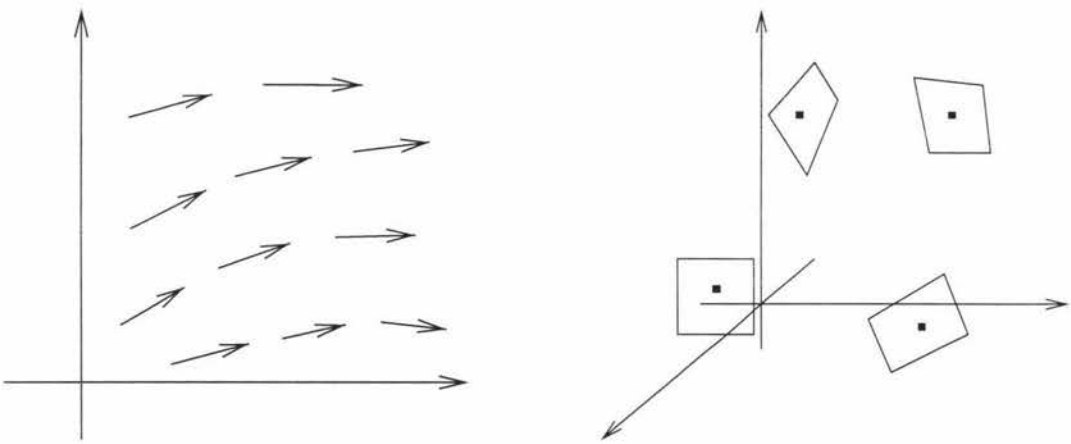
Despite being energy preserving and reversible, simple mechanical systems are not, in general, Hamiltonian [7].

Theorem 6 *The lowest dimension in which a nonholonomic dynamical system can exist is 6.*

Proof:

If we consider a nonholonomic system in $\mathbb{R}^2$ with one constraint, then at each point on the plane there is a only one direction of motion possible. This can be seen in figure 3.1(a). The resultant vector field in the plane is smooth and is thus integrable. Note that if there are two constraints then each point is fixed.

For a nonholonomic system in $\mathbb{R}^3$ with one constraint, each point may move in a plane as in figure 3.1(b). There do exist constraints $A(q)\dot{q} = 0$, $q \in \mathbb{R}^3$, which are not the derivative of some constant $f(q) = 0$; an example is given in the next theorem.



(a) Diagram of a (pseudo) nonholonomic system in $\mathbb{R}^2$. Each point can only move in one direction, so the constraint is really holonomic.

(b) Diagram of a nonholonomic system in $\mathbb{R}^3$. Each point can move in 2-dimensions.

Figure 3.1: Diagrams of nonholonomic systems in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Theorem 7 (Darboux's Theorem for Nonholonomic Constraints) *For every nonholonomic dynamical system of dimension $2n + 1$ with a single constraint, the constraint can be written in some local coordinate system in the "normal form"*

$$\sum_{i=1}^{n} x_i \dot{y}_i + \dot{z} = 0 \tag{3.4}$$

where $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ and z are the local coordinates [1].

Definition 16 *We call a nonholonomic system with a Lagrangian given by equation (3.3) and a single constraint as in equation (3.4) a* contact particle. *(The name comes from an analog with Darboux's theorem for contact structures, equivalent to theorem 7).*

3.4 Examples of Nonholonomic Systems

We now illustrate the application of the Lagrange d'Alembert equations in two cases, the rolling penny and the rattleback.

We will see below that the rolling penny (see figure 3.2) system is symmetric in that it doesn't matter which way the penny is rolling, the dynamics is the same in either direction. The rattleback (see figure 3.4) on the other hand is obviously non-symmetric as can be seen from spinning it, in one direction it spins freely but resists in the other direction. This shows that nonholonomic systems can exhibit both symmetric, reversible behaviour and non-symmetric, reversible behaviour.

The rolling penny system as shown below is too simple in that the system can be solved explicitly, while the rattleback system is too complicated for this thesis as it involves too many dimensions. For this thesis I wanted a class of simple nonholonomic systems that are generally non-integrable. The minimum number of dimensions required for this is 6 as shown in theorem 6 and from Darboux's theorem we can see that the simplest class of nonholonomic systems is the contact particle in $\mathbb{R}^3$. So the contact particle in $\mathbb{R}^3$ is the system I have studied in this thesis.

3.4.1 Rolling Penny [6]

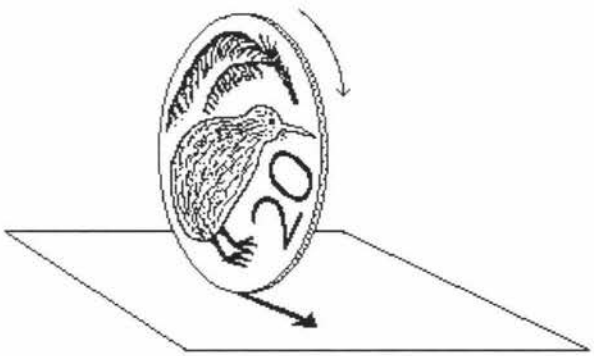


Figure 3.2: Rolling penny: the coin can only move in the direction it is rolling.

The rolling penny problem is as follows: A vertical homogeneous disc of radius r is rolling without slipping on a horizontal plane as shown in figure 3.3. The disc can only roll in the direction of the plane parallel to the plane of the disc, but the plane of the disc can be rotated

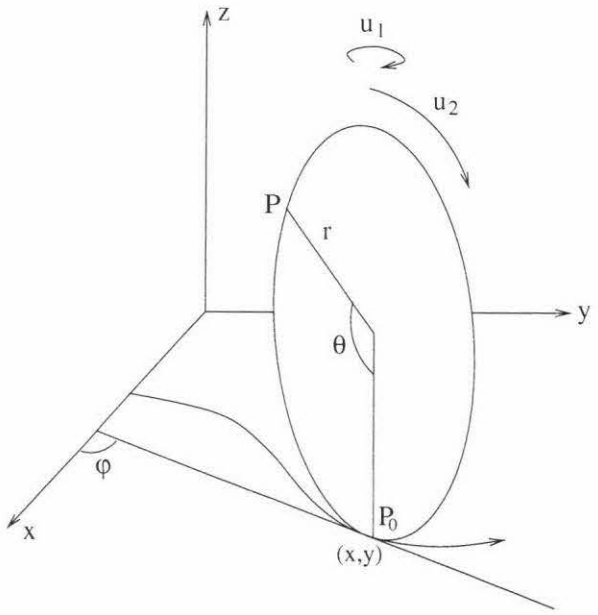


Figure 3.3: Diagrammatic representation of the rolling penny.

so as to roll in a different direction by an external control force u_1, giving steerage. Also, an external control force u_2 gives control over the amount of roll torque.

The configuration space of the rolling disc is $Q = \mathbb{R}^2 \times S^1 \times S^1$ with co-ordinates $q = (x, y, \theta, \varphi)$ where θ is the rotation angle and φ is the orientation of the disc. The Lagrangian for this system is

$$L(x, y, \theta, \varphi, \dot{x}, \dot{y}, \dot{\theta}, \dot{\varphi}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\varphi}^2$$

where m is the mass of the disc, I is the moment of inertia perpendicular to the plane of the disc and J is the moment of inertia parallel to the plane of the disc.

The nonholonomic constraints are

$$\begin{aligned} \dot{x} &= r(\cos\varphi)\dot{\theta} \\ \dot{y} &= r(\sin\varphi)\dot{\theta} \end{aligned}$$

and are such that the point P_0 has zero velocity when it is in contact with the plane.

The Lagrange d'Alembert equations give the equations of motion to be

$$\begin{aligned} m\ddot{x} &= \lambda_1 \\ m\ddot{y} &= \lambda_2 \\ I\ddot{\theta} &= -\lambda_1 r\cos(\varphi) - \lambda_2 r\sin(\varphi) \\ J\ddot{\varphi} &= 0 \end{aligned}$$

with the constraints

$$\begin{aligned} \dot{x} &= r\cos(\varphi)\dot{\theta} \\ \dot{y} &= r\sin(\varphi)\dot{\theta}. \end{aligned}$$

From the first three equations of motion we get

$$I\ddot{\theta} = -m\ddot{x}r\cos(\varphi) - m\ddot{y}r\sin(\varphi)$$

but differentiating the constraints gives

$$\begin{aligned}\ddot{x} &= r\cos(\varphi)\ddot{\theta} - r\sin(\varphi)\dot{\varphi}\dot{\theta}\\ \ddot{y} &= r\sin(\varphi)\ddot{\theta} + r\cos(\varphi)\dot{\varphi}\dot{\theta}\end{aligned}$$

and so

$$\begin{aligned}I\ddot{\theta} &= -m(r\cos(\varphi)\ddot{\theta} - r\sin(\varphi)\dot{\varphi}\dot{\theta})r\cos(\varphi) - m(r\sin(\varphi)\ddot{\theta} + r\cos(\varphi)\dot{\varphi}\dot{\theta})r\sin(\varphi)\\ &= -mr^2\ddot{\theta}.\end{aligned}$$

Thus, with the control forces, the equations of the dynamics are

$$\begin{aligned}(I + mr^2)\ddot{\theta} &= u_1\\ J\ddot{\varphi} &= u_2\\ \dot{x} &= r(\cos\varphi)\dot{\theta}\\ \dot{y} &= r(\sin\varphi)\dot{\theta}.\end{aligned}$$

The free equations with $u_1 = u_2 = 0$ can be integrated directly to give

$$\begin{aligned}\theta &= \Omega t + \theta_0\\ \varphi &= \omega t + \varphi_0\\ x &= \tfrac{\Omega}{\omega}R\sin(\omega t + \varphi_0) + x_0\\ y &= -\tfrac{\Omega}{\omega}R\cos(\omega t + \varphi_0) + y_0.\end{aligned}$$

3.4.2 Rattleback

The rattleback is a convex non-symmetric rigid body that is rolling without sliding on a horizontal plane. It is also known as a wobblestone and a celt. If spun in a particular direction the rattleback spins freely but if spun in the opposite direction the rattleback resists, rocks back and forth approximately along the long axis before reversing the direction of the spin, hence the name rattleback. This corresponds to a stable and an unstable orbit of the system. This type of system is referred to by Nordmark and Essén [28] as a 'chiral top'. There are two essential conditions to the dynamics of the rattleback,

- The rattleback rolls without slipping and without resistance, thus energy is conserved.
- The principal directions of the curvature of the contact surface are not aligned with the principal axes of the inertia tensor.

Despite conservation of energy, the system is non-Hamiltonian as the nonholonomic rolling constraint destroys the Hamiltonian structure of the system.

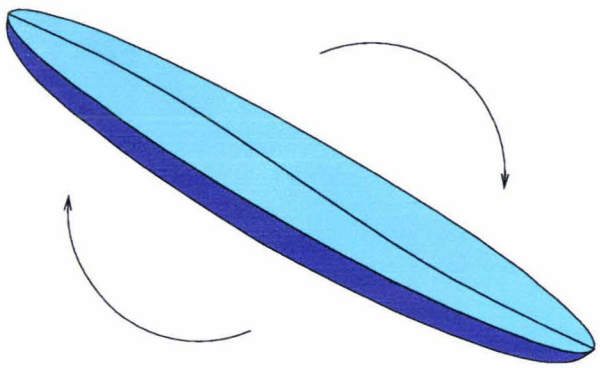


Figure 3.4: Rattleback: Spinning in one direction is stable, spinning in the other direction is unstable.

The shape of the 'hull' of the boat shaped rattleback can be expressed as:

$$z = a\left[1 - \frac{1}{2}p\left(\frac{x}{a}\right)^2 - q\frac{xy}{a^2} - \frac{1}{2}s\left(\frac{y}{a}\right)^2\right]$$

where a is the rest distance from the centre of mass to the point of contact with the horizontal surface; there are no first order terms because the rest position is at $x = y = 0$. It can be shown that this surface is convex with a stable rest position if

$$0 < p < 1,\ 0 < s < 1,\ ps > q^2,\ (1-p)(1-q) > q^2$$

and chirality occurs unless $q = 0$ [11].

If we define the Euler angles θ, ϕ, φ using the principal axis body frame relative to an inertial reference frame and use the horizontal coordinates x, y of the center of mass, then the configuration space of the rattleback is $Q = SO(3) \times \mathbb{R}^2$ with coordinates $q = (\theta, \varphi, \psi, x, y)$. The Lagrangian for the system is [6]

$$\begin{aligned} L &= \tfrac{1}{2}[A\cos^2\psi + B\sin^2\psi + m(\gamma_1\cos\theta - \zeta\sin\theta)^2]\dot{\theta}^2 \\ &\quad + \tfrac{1}{2}[(A\sin^2\phi + B\cos^2\phi)\sin^2\theta + C\cos^2\theta]\dot{\varphi}^2 \\ &\quad + \tfrac{1}{2}(C + m\gamma_2^2\sin s^2\theta)\dot{\psi}^2 + \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) \\ &\quad + m(\gamma_1\cos\theta - \zeta\sin\theta)\gamma_2(\sin\theta)\dot{\theta}\dot{\psi} + (A - B)(\sin\theta)(\sin\psi)(\cos\psi)\dot{\theta}\dot{\varphi} \\ &\quad + C(\cos\theta)\dot{\varphi}\dot{\psi} + mg(\gamma_1\sin\theta + \zeta\cos\theta) \end{aligned}$$

where

$$\begin{aligned} A, B, C &= \text{principal moments of inertia of the body,} \\ m &= \text{total mass of the body,} \\ (\xi, \eta, \zeta) &= \text{coordinates of the point of contact} \\ &\quad \text{relative to the body frame of reference,} \\ \gamma_1 &= \xi\sin\psi + \eta\cos\psi, \\ \gamma_2 &= \xi\cos\psi - \eta\sin\psi. \end{aligned}$$

The constraints for this system are

$$\begin{aligned} \dot{x} &= \alpha_1\dot{\theta} + \alpha_2\dot{\psi} + \alpha_3\dot{\varphi} \\ \dot{y} &= \beta_1\dot{\theta} + \beta_2\dot{\psi} + \beta_3\dot{\varphi} \end{aligned}$$

where

$$\begin{aligned} \alpha_1 &= -(\gamma_1 \sin\theta + \zeta\cos\theta)\sin\varphi \\ \alpha_2 &= \gamma_2\cos\theta\sin\varphi + \gamma_1\cos\varphi \\ \alpha_3 &= \gamma_2\sin\varphi + (\gamma_1\cos\theta - \zeta\sin\theta)\cos\varphi \\ \beta_k &= -\tfrac{\partial\alpha_k}{\partial\varphi}, \qquad k = 1,2,3 \end{aligned}$$

3.4.3 Contact particle in $\mathbb{R}^3$

From theorems 6 and 7 it can be seen that the contact particle in $\mathbb{R}^3$ is the simplest nonholonomic system possible. The constraint for this class of systems can be written as

$$p_1 + q_2 p_3 = 0. \tag{3.5}$$

If we have the Hamiltonian

$$H = \frac{1}{2}\|p\|^2 + V(q) \tag{3.6}$$

then we get the equations of motion to be:

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -\nabla V(q) + \lambda \begin{pmatrix} 1 \\ 0 \\ q_2 \end{pmatrix}. \end{aligned} \tag{3.7}$$

For a spherically symmetric potential energy $V = \frac{1}{2}\|q\|^2$, which gives us a Hamiltonian

$$H = \tfrac{1}{2}\|p\|^2 + \tfrac{1}{2}\|q\|^2. \tag{3.8}$$

The equations of motion are:

$$\begin{aligned} p_1 + q_2 p_3 &= 0 \\ \dot{q}_1 &= p_1 \\ \dot{q}_2 &= p_2 \\ \dot{q}_3 &= p_3 \\ \dot{p}_1 &= -q_1 + \lambda \\ \dot{p}_2 &= -q_2 \\ \dot{p}_3 &= -q_3 + \lambda q_2 \end{aligned} \tag{3.9}$$

Although nonlinear, we can still describe the behaviour of this system, as I now show. Using $\dot{p}_1 = -q_1 + \lambda$ together with the constraint $p_1 = -q_2 p_3$, we find that

$$\begin{aligned} & -\dot{q}_2 p_3 - q_2\dot{p}_3 = -q_1 + \lambda \\ \Rightarrow\quad & -p_2 p_3 - q_2(-q_3 + \lambda q_2) = -q_1 + \lambda \\ \Rightarrow\quad & \lambda(1 + q_2^2) = q_1 + q_2 q_3 - p_2 p_3 \\ \Rightarrow\quad & \lambda = \tfrac{q_1 + q_2 q_3 - p_2 p_3}{1 + q_2^2} \end{aligned}$$

From the above equations of motion it can be seen that the q_2, p_2 variables decouple from the system and form a simple 1-dimensional harmonic oscillator system.

$$\begin{aligned} I &= \tfrac{1}{2}(p_2^2 + q_2^2) \\ \dot{q}_2 &= p_2 \\ \dot{p}_2 &= -q_2 \end{aligned}$$

Also, we can use the constraint to lower the dimension of the system to just 3 coordinates, q_1, p_3, q_3. This reduced system is then integrable, as I now show:

First, by choosing the initial time suitably, we can write the solution of the decoupled q_2, p_2 system as

$$q_2 = A\sin(t), \qquad p_2 = A\cos(t)$$

where $I = \frac{1}{2}A^2 = \frac{1}{2}(q_2^2 + p_2^2)$ is the energy in the q_2, p_2 subsystem. Then, because of the constraint we can drop the p_1 equation, which gives

$$\begin{aligned} H - I &= \tfrac{1}{2}((-q_2 p_3)^2 + p_3^2) + \tfrac{1}{2}(q_1^2 + q_3^2) \\ \lambda &= (1 + A^2\sin^2(t))^{-1}[q_1 + A\sin(t)q_3 - A\cos(t)p_3] \end{aligned}$$

$$\begin{aligned} \dot{q}_1 &= -A\sin(t)p_3 \\ \dot{p}_3 &= -q_3 + \lambda A\sin(t) \\ &= (1 + A^2\sin^2(t))^{-1}[-(1 + A^2\sin^2(t))q_3 + (q_1 + A\sin(t)q_3 - A\cos(t)p_3)A\sin(t)] \\ &= (1 + A^2\sin^2(t))^{-1}[A\sin(t)q_1 - q_3 - A^2\sin(t)\cos(t)p_3] \\ \dot{q}_3 &= p_3 \end{aligned}$$

Now letting $\tilde{p}_3 = (1 + A^2\sin^2(t))^{\frac{1}{2}}p_3$ gives

$$\begin{aligned} &\dot{\tilde{p}}_3(1 + A^2\sin^2(t))^{-\frac{1}{2}} - \tfrac{1}{2}\tilde{p}_3(1 + A^2\sin^2(t))^{-\frac{3}{2}} \times 2A^2\sin(t)\cos(t) \\ &\qquad = (1 + A^2\sin^2(t))^{-1}[A\sin(t)q_1 - q_3 - A^2\sin(t)\cos(t)p_3], \end{aligned}$$

which, after rearranging gives

$$\dot{\tilde{p}}_3 = (1 + A^2\sin^2(t))^{-\frac{1}{2}}[A\sin(t)q_1 - q_3].$$

We can then write the system as

$$\begin{pmatrix} \dot{q}_1 \\ \dot{\tilde{p}}_3 \\ \dot{q}_3 \end{pmatrix} = \begin{pmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & -\beta \\ 0 & \beta & 0 \end{pmatrix} \begin{pmatrix} q_1 \\ \tilde{p}_3 \\ q_3 \end{pmatrix} \tag{3.10}$$

where

$$\begin{aligned} \alpha &= -A\sin(t)(1 + A^2\sin^2(t))^{-\frac{1}{2}}, \\ \beta &= (1 + A^2\sin^2(t))^{-\frac{1}{2}}. \end{aligned}$$

The above matrix is an element of the Lie algebra $so(3)$ and thus the solution of equation (3.10) has the form

$$x(2\pi) = M(A)x(0)$$

for some $M(A) \in SO(3)$. That is, the time-2π flow is a rotation, so the system has an extra integral defined by the axis of rotation. The full system on the constraint manifold $\mathbb{R}^5$ has 3 integrals H, I and the axis of rotation; all orbits are either fixed points, periodic points or 2-tori.

We cannot find the rotation matrix explicitly but we can approximate it for small A using the Magnus expansion [19], as follows. Representing the solution of a time-dependent linear ODE such as equation (3.10) in the form

$$x(t) = \exp(\Theta(t))x(0),$$

Θ can be obtained to arbitrary order by iteratively applying

$$\Theta^{[m+1]}(t) = \sum_{k=0}^{\infty} \frac{B_k}{k!} \int_0^t ad^k_{\Theta^{[m]}(s)} B(s) \mathrm{d}s,$$

where $\Theta^{[0]} = 0$, B_k are the Bernoulli numbers, $1, -\frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \frac{5}{66}, \ldots$, and $B(s)$ is an element of a Lie algebra, e.g. the matrix in equation 3.10. Also, $ad^k_{\Theta^{[m]}(s)} B(s)$ is the k-th adjoint (or Lie bracket) of $\Theta^{[m]}(s)$ with $B(s)$, i.e. $[\Theta^{[m]}(s), [\Theta^{[m]}(s), \ldots, [\Theta^{[m]}(s), B(s)]\ldots]]$, with the 0-th adjoint being just $B(s)$.

Rewriting equation 3.10 as

$$\dot{X} = \xi(t)X \qquad X \in \mathbb{R}^3,\ \xi(t) \in so(3), \tag{3.11}$$

we can remove the constant term of $\xi(t)$ to improve the convergence of the Magnus expansion by writing $\xi(t)$ as

$$\xi(t) = \xi_0 + A\xi_1 + A^2\xi_2 + \ldots$$

where

$$\xi_0 = \xi(0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Then we write

$$\begin{aligned} X &= e^{t\xi_0} X_1 \\ \Rightarrow \quad \dot{X} &= \xi_0 e^{t\xi_0} X_1 + e^{t\xi_0} \dot{X}_1, \end{aligned}$$

and substituting in equation 3.11 gives

$$\begin{aligned} \xi(t)X &= \xi_0 X + e^{t\xi_0} \dot{X}_1 \\ \Rightarrow \quad \dot{X}_1 &= e^{-t\xi_0}(\xi(t) - \xi_0) e^{t\xi_0} X_1 \\ &= B(t) X_1, \end{aligned}$$

where

$$B(t) = \begin{pmatrix} 0 & \alpha\cos(t) & -\alpha\sin(t) \\ -\alpha\cos(t) & 0 & 1-\beta \\ \alpha\sin(t) & \beta - 1 & 0 \end{pmatrix},$$

and α and β are as given above. Now we have $B(t) = \mathcal{O}(A)$, $\Theta(t) = \mathcal{O}(A)$, and $ad^k_\Theta B = \mathcal{O}(A^{k+1})$, which allows us to truncate the infinite series in equation (3.4.3) to the desired order.

Using the matrix $B(t)$ and the code in Appendices A.2.1 and A.2.2, I was able to compute (in MATLAB) the rotation matrix to be:

$$\Theta^{[8]} = \pi \begin{bmatrix} 0 & 0 & \gamma \\ 0 & 0 & \delta \\ -\gamma & -\delta & 0 \end{bmatrix}$$

where

$$\gamma = A - \tfrac{17}{64}A^3 + \tfrac{293}{2048}A^5 - \tfrac{100327}{1048576}A^7$$
$$\delta = \tfrac{1}{8}A^2 - \tfrac{11}{128}A^4 + \tfrac{4127}{65536}A^6 - \tfrac{2469695}{54331648}A^8$$

to order 8 in A. As can be seen from γ and δ the magnitude of the coefficients of A are slowly decreasing and the series will converge for $|A| \lessapprox 1$.

Thus for small A the axis of rotation is:

$$\begin{bmatrix} \delta \\ -\gamma \\ 0 \end{bmatrix}$$

and the amount of rotation is:

$$\pi(\gamma^2 + \delta^2)^{\frac{1}{2}}$$

This was confirmed by direct numerical solution as shown in figure 3.5.

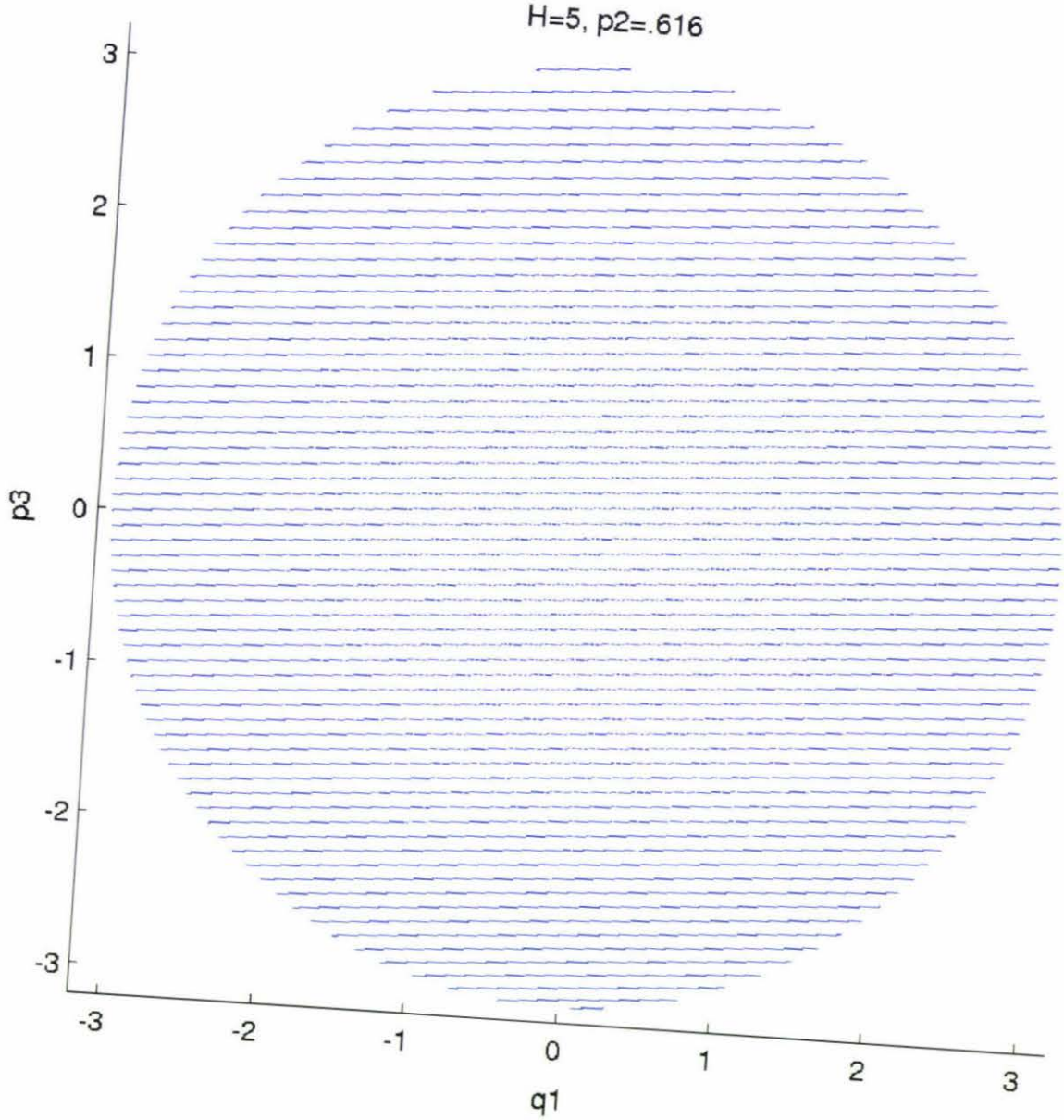


Figure 3.5: Circles on the sphere with an energy of $H = 5$, approximately 0.616 of which was in the q_2, p_2 subsystem. Confirming the axis of rotation of the invariant sphere, the slope of the axis of rotation here is approximately $\frac{-\gamma}{\delta} = -15$.

Chapter 4

Numerical Studies of Nonholonomic Systems in 6 Dimensions

4.1 Strategy

4.1.1 The Approach

The main goal of this thesis is to find evidence of dissipative behaviour in nonholonomic dynamical systems. In section 3.4 I have shown that theorems 6 and 7 mean that the contact particle in $\mathbb{R}^3$ is a possible system.

To define what is meant by dissipative behaviour in a system we need to be able to describe the orbits of the system in forwards and backwards time; to do this it is useful to have the notion of ω- and α-limit sets.

Definition 17 (Limit set) *[32] A point y is an* ω-limit point *of x for the flow ϕ_t if there exists a sequence of times t_k going to infinity such that $\lim_{k\to\infty} d(\phi_{t_k}(x), y) = 0$. The* ω-limit set *of x is the set of all such points y and is denoted by $\omega(x)$. An* α-limit set *is defined the same way but with t_k going to negative infinity.*

Some properties of limit sets of flows, which are also valid for maps, are given in the next theorem:

Theorem 8 *[32] Let ϕ_t be a flow on a metric space M.*

(a) The ω-limit set can be represented in terms of the forward orbit as follows:

$$\omega(x) = \bigcap_{T\geq 0} cl \bigcup_{t\geq T} \{\phi_t(x)\}.$$

(b) If $\phi_t(x) = y$ for a real number t, then $\omega(x) = \omega(y)$ and $\alpha(x) = \alpha(y)$.

(c) The limit sets, $\omega(x)$ and $\alpha(x)$, are closed and both positively and negatively invariant, i.e. they contain complete orbits.

(d) If $y \in \omega(x)$, then $\omega(y) \subset \omega(x)$ and $\alpha(y) \subset \alpha(x)$.

Thus we can define a *dissipative orbit* as one in which the ω- and α-limit sets occupy different regions of the phase space. A region of a system exhibits *dissipative behaviour* if that region contains dissipative orbits.

As mentioned in section 2.3, reversible systems are non-dissipative near symmetric orbits and (possibly) dissipative away from symmetric orbits. This can be explained by the following:

Theorem 9 *Consider a reversible dynamical system with reversing symmetry R. The ω- and α-limit sets of a point x are related by*

$$\omega \circ R(x) = R \circ \alpha(x).$$

If the orbit containing the point x is symmetric with respect to R, then

$$\omega(x) = \alpha(x).$$

Proof:

Consider the flow $\phi_t(x)$, then from the definition of reversibility we find

$$\begin{aligned} & \phi_t(x) = R \circ \phi_{-t} \circ R^{-1}(x) \\ \Rightarrow \quad & \omega(x) = R \circ \alpha \circ R^{-1}(x) \\ \Rightarrow \quad & \omega \circ R(x) = R \circ \alpha(x). \end{aligned}$$

If the orbit is symmetric then

$$\begin{aligned} & \omega(x) = \omega \circ R(x) \text{ and } \alpha(x) = R \circ \alpha(x) \\ \Rightarrow \quad & \omega(x) = \alpha(x). \end{aligned}$$

Thus, symmetric orbits cannot be dissipative. If, however, the orbit is not symmetric then it is possible to have $\omega(x) \neq \alpha(x)$, giving dissipative orbits.

For volume preserving systems, such as Hamiltonian systems, the ω- and α-limit sets of almost all points in the system must be the same due to Poincaré's recurrence theorem:

Theorem 10 (Poincaré Recurrence Theorem[32]) *Assume μ is a finite measure on a space M, $\mu(M) = \int_\mu M < \infty$. Assume $f : M \mapsto M$ is a one to one map which preserves the measure μ, i.e. $\mu(f^{-1}(A)) = \mu(A)$ for every measurable set A. Assume S is a measurable set. Then, for almost all $x \in S$, $f^j(x)$ returns to S infinitely often.*

In addition, for the integrable nonholonomic systems I have studied, all the orbits (even the non-symmetric orbits) are quasi-periodic and hence have $\omega(x) = \alpha(x)$. We therefore want to study non-integrable nonholonomic systems with many non-symmetric orbits. These can be found as perturbations of simple integrable systems with many non-symmetric orbits. The integrable base system that I started with was initially a harmonic oscillator system with a spherically symmetric potential, but later I changed this to a system with a free particle in the q_3 direction.

We solved each integrable base system analytically by the method in section 3.4.3 with some help from the Magnus expansion for the spherically symmetric potential. I have also written code that I used to numerically integrate the systems, which will be elaborated upon in the next sections. Both methods give the same result which I used as confirmation that my code works.

4.1.2 Discretised Contact Particle

As stated in section 1.1, a simple mechanical nonholonomic system in $\mathbb{R}^3$, with a mass matrix $\mu(q) = Id$, can be written as

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -\nabla V(q) + A^T(q)\lambda \\ A(q)p &= 0. \end{aligned}$$

For the contact particle, we have $A(q) = (1, 0, q_2)$, giving the system in equation (3.7). This system can be solved analytically for some $V(q)$, as was shown in section 3.4.3, but in general, a numerical integrator must be used. To numerically integrate this system on a computer I used the following integrator which takes the current values of (q_i, p_i), satisfying $A(q_i)p_i = 0$, and returns the next values (q_{i+1}, p_{i+1}) by the following process. Let Δt be the time step, then

$$\begin{aligned} \tilde{q} &= q_i + \tfrac{1}{2}p_i\Delta t \\ \tilde{p} &= p_i - \nabla V(\tilde{q})\Delta t \\ p_{i+1} &= \tilde{p} + A^T(\tilde{q})\lambda \\ q_{i+1} &= \tilde{q} + \tfrac{1}{2}p_{i+1}\Delta t \end{aligned}$$

where

$$p_i = \begin{pmatrix} p_{1,i} \\ p_{2,i} \\ p_{3,i} \end{pmatrix}, \; q_i = \begin{pmatrix} q_{1,i} \\ q_{2,i} \\ q_{3,i} \end{pmatrix},$$

and $\tilde{q}$, $\tilde{p}$ are temporary vectors. The value of λ can be derived as follows. We require the constraint, $A(q_{i+1})p_{i+1} = 0$, to be satisfied after each step. Then substituting the temporary variables gives

$$A(\tilde{q} + \frac{1}{2}\Delta t(\tilde{p} + A^T(\tilde{q})\lambda))(\tilde{p} + A^T(\tilde{q})\lambda) = 0. \tag{4.1}$$

In general this is non-linear in λ. If A is linear then equation (4.1) is at most quadratic in λ. This method is semi-explicit (explicit in q and p and implicit in λ), reversible and second order in the time step, Δt.

However, for some A the above method may become explicit. For example, for the contact particle $A = (1, 0, q_2)$, which is not only linear, but because of the 0 value, it makes equation (4.1) linear in λ. That is,

$$\begin{array}{rrcl} & (\ \tilde{p}_1 + \lambda) + (q_{2,i+1})(\tilde{p}_3 + \lambda\tilde{q}_2) & = & 0 \\ \Rightarrow & \lambda(1 + q_{2,i+1}\tilde{q}_2) & = & -(\tilde{p}_1 + q_{2,i+1}\tilde{p}_3) \\ \Rightarrow & \lambda & = & -\frac{\tilde{p}_1 + q_{2,i+1}\tilde{p}_3}{1+q_{2,i+1}\tilde{q}_2} \end{array}$$

The orbits that result from integrating the above system are in six dimensions. Since it is nigh impossible to visualise these, I have reduced the six dimensional manifold of the system, $M = \mathbb{R}^6$, to three dimensions. This is done via two observations and a Poincaré section.

Firstly, the system preserves the energy $H(q,p) = \frac{1}{2}p^2 + V(q)$, thus we can lower the dimension by one. Since, in this case, the unperturbed system has the q_2, p_2 subsystem decoupling, q_2 or p_2 are the obvious coordinates to drop.

Secondly, the constraint $p_1 = -q_2 p_3$ lowers the dimension by another one as we can write any function of p_1 as a function of q_2 and p_3. The intersection of the two 5-dimensional submanifolds of M, $\{(q,p) : H(q,p) = h_0\}$ and $\{(q,p) : p_1 + q_2 p_3 = 0\}$, gives a 4-dimensional submanifold of M.

Thirdly, I have taken a Poincaré section at $q_2 = 0, \dot{q}_2 = p_2 > 0$ as this forces $p_1 = 0$ and we can drop p_2 due to the first observation. Thus the resulting 3-dimensional Poincaré section is

$$\begin{array}{rcl} \Sigma & = & \{(q,p) : H(q,p) = h_0, p_1 + q_2 p_3 = 0, q_2 = 0, \dot{q}_2 = p_2 > 0\} \\ & = & \{(q,p) : H(q,p) = h_0, p_1 = 0, q_2 = 0, p_2 > 0\}. \end{array}$$

Because of the condition $p_2 > 0$, we can take coordinates (q_1, p_3, q_3) on Σ and obtain a unique solution for p_2 in $H(q,p) = h_0$.

The Poincaré map is implemented as a linear interpolation between the steps surrounding the section. That is, when the previous step and the current step straddle the chosen section in the right direction, a straight line is constructed between the two points and the point where that line crosses the section is taken. The error in this method is proportional to $(\Delta t)^2$ and is a source of numerical inaccuracy.

In order to observe the 3-dimensional Poincaré map, I used MATLAB to plot the forwards and backwards orbits, that were returned from the integrator code, in different colours. This allows for differing ω- and α-limit sets to be seen immediately.

4.2 Harmonic Oscillator

Initially, I studied the Poincaré map, $\phi : \Sigma \mapsto \Sigma$, of a harmonic oscillator system in three dimensions. This has a potential that is spherically symmetric, $V(q) = \frac{1}{2}\|q\|^2$, with the dynamics described in section 3.4.3 by the system of equations (3.9).

As mentioned earlier, the q_2, p_2 system decouples and the Poincaré section at $q_2 = 0$ forces $p_1 = 0$ leaving the coordinates (q_1, p_3, q_3) which can be solved via the Magnus expansion in

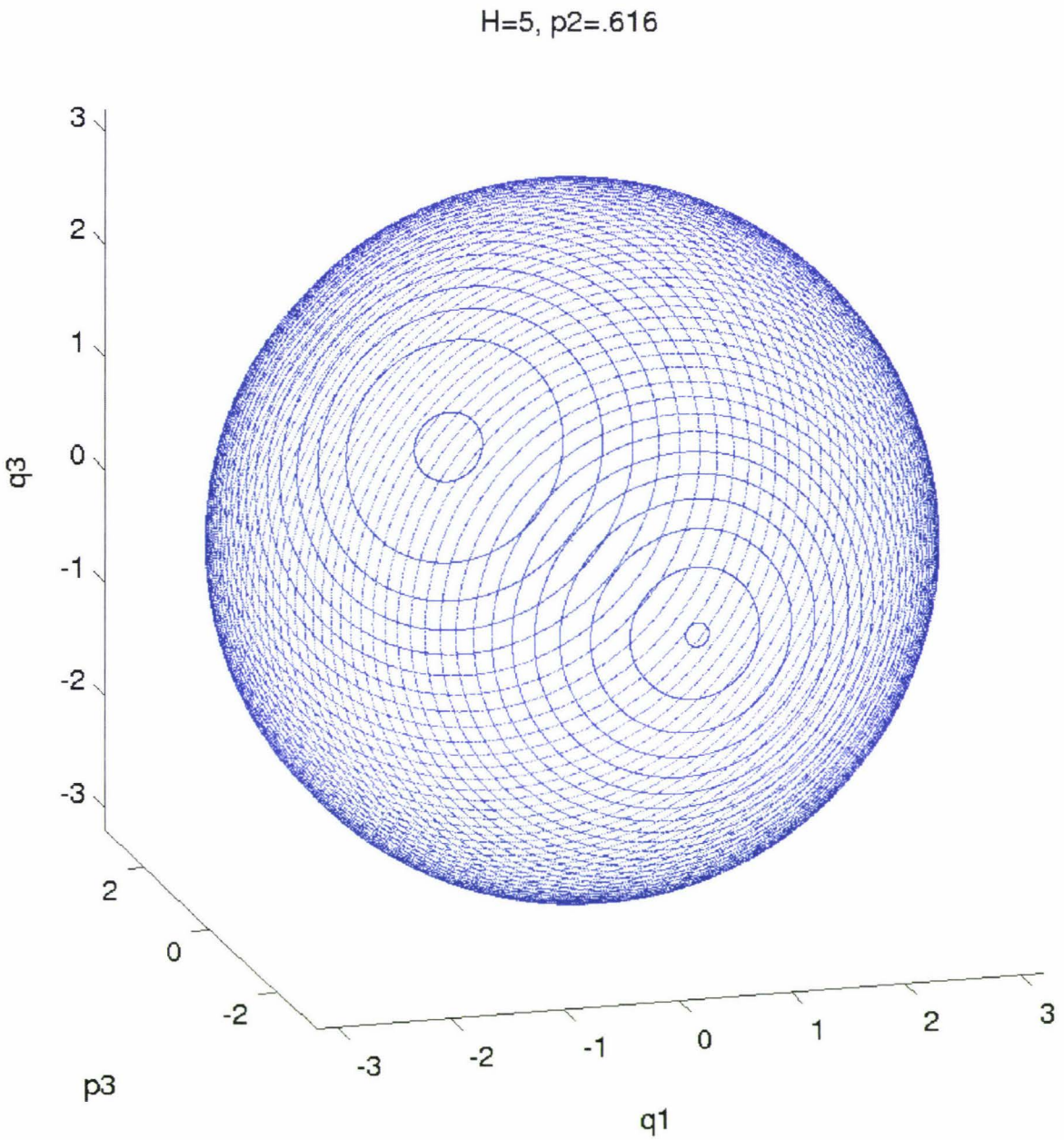


Figure 4.1: Circular orbits on the sphere generated by the unperturbed harmonic oscillator with energy $H = 5$, approximately 0.616 of which is in the q_2, p_2 subsystem. Note: this is the same set of orbits as those of figure 3.5 and the axis of rotation is nearly parallel to the p_3 axis.

Appendix A.2.1 to give a rotation about a fixed axis. Thus all the orbits of the Poincaré map are circles on the sphere centred on the rotation axis as can be seen in figure 4.1.

This system has at least four symmetries and reversing symmetries, the symmetries are:

$$\begin{array}{lrcll} 1 & (q_1,q_2,q_3,p_1,p_2,p_3) & \mapsto & (q_1,q_2,q_3,p_1,p_2,p_3) & (I) \\ 2 & (q_1,q_2,q_3,p_1,p_2,p_3) & \mapsto & (-q_1,q_2,-q_3,-p_1,p_2,-p_3) & \\ 3 & (q_1,q_2,q_3,p_1,p_2,p_3) & \mapsto & (-q_1,-q_2,q_3,-p_1,-p_2,p_3) & \\ 4 & (q_1,q_2,q_3,p_1,p_2,p_3) & \mapsto & (q_1,-q_2,-q_3,p_1,-p_2,-p_3) & \end{array}$$

and the reversing symmetries are:

$$\begin{array}{lrcll} 1 & (q_1,q_2,q_3,p_1,p_2,p_3) & \mapsto & (q_1,q_2,q_3,-p_1,-p_2,-p_3) & (R) \\ 2 & (q_1,q_2,q_3,p_1,p_2,p_3) & \mapsto & (-q_1,q_2,-q_3,p_1,-p_2,p_3) & \\ 3 & (q_1,q_2,q_3,p_1,p_2,p_3) & \mapsto & (-q_1,-q_2,q_3,p_1,p_2,-p_3) & \\ 4 & (q_1,q_2,q_3,p_1,p_2,p_3) & \mapsto & (q_1,-q_2,-q_3,-p_1,p_2,p_3) & \end{array}$$

In my first attempt to integrate this system I implemented the integrator of section 4.1.2 in a MATLAB m-file, but the run times were long even for small numbers of returned points. So I rewrote the integrator in C using MATLAB's "mex" interface and managed to get about a 100,000-fold increase in speed, allowing many points to be returned from a short run. The C code that I have written to perform the integration and the Poincaré section can be found in Appendix A.1.1.

For very long runs with large numbers of points in the Poincaré map the arrays returned from the integrator code are very large. Due to memory restrictions a significant alteration of the integrator code was needed for these situations. Instead of returning every point, the phase space was divided up into a grid in which each element recorded the number of "hits", giving a density map of the phase space and requiring only a fixed amount of memory allowing for arbitrarily long runs. This code is given in Appendix A.1.2.

4.2.1 Perturbations of the Harmonic Oscillator

To the initial potential, I then added various perturbations, $\epsilon V_1(q)$, giving

$$V = \frac{1}{2}\|q\|^2 + \epsilon V_1(q)$$

and studied the Poincaré map, $\phi : \Sigma \mapsto \Sigma$, as ϵ increases, to observe the breakdown of integrability and the possible onset of dissipative behaviour.

When considering perturbations of the system I wanted the perturbation to be as simple as possible.

"Everything should be made as simple as possible, but not simpler." – Albert Einstein.

Another issue that I had to take into consideration was the size of ϵ in the perturbation. The first part is a usability issue. If ϵ is not large enough then the system is not perturbed

enough from the base system and the integrator code needs to be run with a very small time-step to be sure that differences in the orbits occur as a result of the dynamics and not from a lack of accuracy. This is expensive in terms of processor time and is undesirable.

The second part is to do with chaos. If ϵ is too large then most of the orbits become chaotic or are at least strongly driven by the perturbation. This causes most of the orbits which previously did not cross the fixed set, $\mathrm{Fix}(R)$, to cross the fixed set and hence be symmetric.

Generally, for this system, I used a value between 0.01 and 0.5 for ϵ, with the smaller values not having much effect on the system and the larger values causing too much chaos.

One of the perturbations that I added was $\epsilon q_1 q_3^2$ which deforms the sphere in the manner shown in figures 4.2(a)-4.2(i) as ϵ is increased.

With this perturbation the system is:

$$\begin{aligned} \dot{q} &= p \\ \dot{p_1} &= -q_1 + \lambda - \epsilon q_3^2 \\ \dot{p_2} &= -q_2 \\ \dot{p_3} &= -q_3 + \lambda q_2 - 2\epsilon q_1 q_3 \end{aligned}$$

with the constraint

$$p_1 + q_2 p_3 = 0.$$

This perturbation breaks the 2nd and 3rd symmetries and reversing symmetries, but retains the 1st and 4th symmetries and reversing symmetries. This 4th reversing symmetry, I eventually realised, was actually causing all the orbits to be symmetric in this case. The fixed set of the 4th reversing symmetry is $\mathrm{Fix}(R_4) = \{(q_1, p_2, p_3) : q_2 = q_3 = p_1 = 0\}$, which is in the Poincaré section. The axis of rotation of the unperturbed system is also contained in the fixed set, so almost all of the orbits intersect the fixed set. This can be seen in figure 4.3. Thus almost all of the orbits are symmetric with respect to this reversing symmetry and cannot be dissipative.

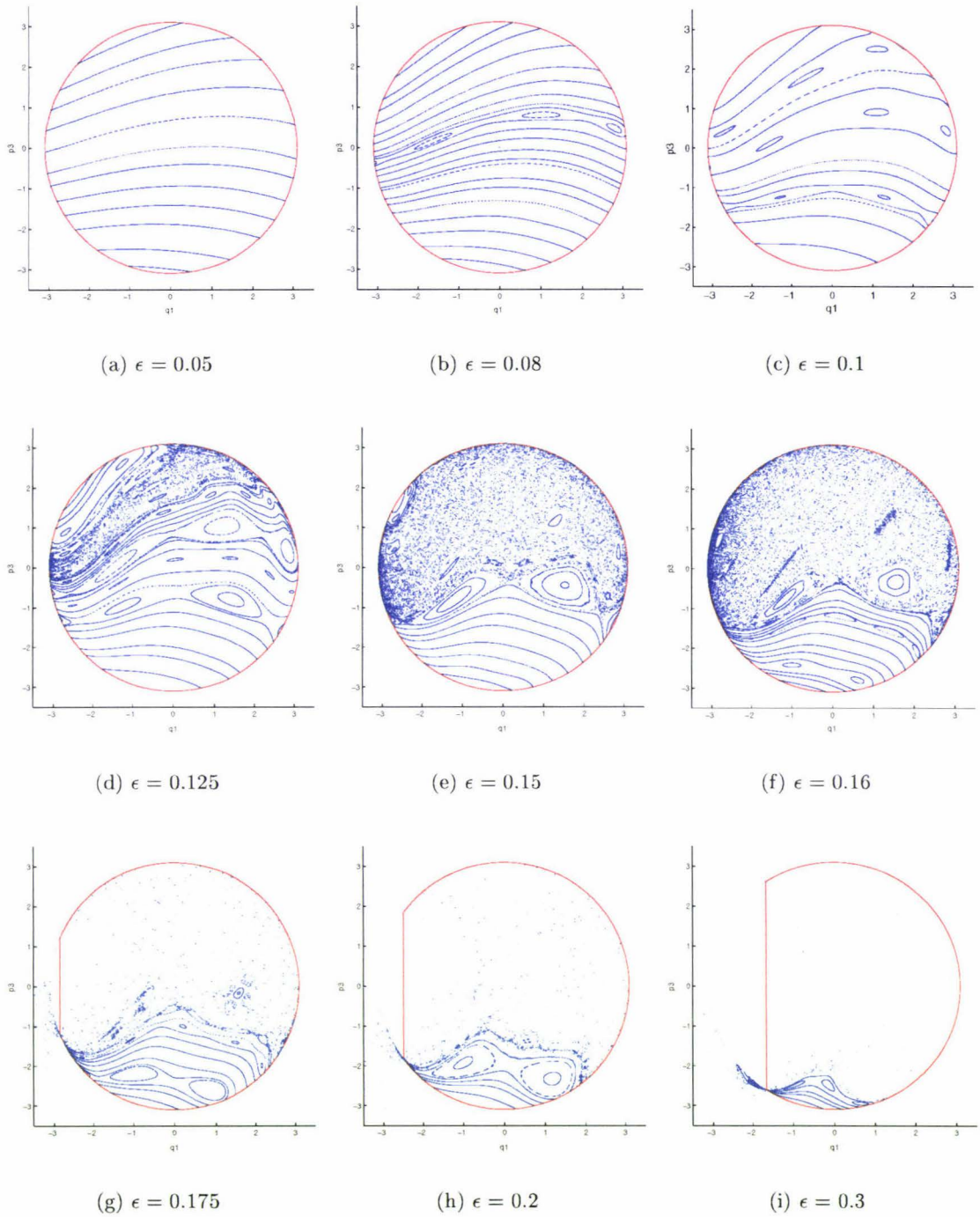
For this perturbation, and the unperturbed system, all the orbits in the Poincaré section are actually in 2 dimensions. They are contained on the surface of the (perturbed) sphere (or non-compact surface, for large ϵ). This is due to the q_2, p_2 subsystem decoupling, so no energy can be transferred from the p_2 coordinate to the q_1, p_3, q_3 coordinates of the Poincaré section. This could possibly account for the lack of dissipative behaviour seen in this system.

The next perturbation that I considered was $\epsilon q_2 q_3^2$, which gives the system

$$\begin{aligned} \dot{q} &= p \\ \dot{p_1} &= -q_1 + \lambda \\ \dot{p_2} &= -q_2 - \epsilon q_3^2 \\ \dot{p_3} &= -q_3 + \lambda q_2 - 2\epsilon q_2 q_3 \end{aligned}$$

with the constraint

$$p_1 + q_2 p_3 = 0.$$



(a) $\epsilon = 0.05$ (b) $\epsilon = 0.08$ (c) $\epsilon = 0.1$

(d) $\epsilon = 0.125$ (e) $\epsilon = 0.15$ (f) $\epsilon = 0.16$

(g) $\epsilon = 0.175$ (h) $\epsilon = 0.2$ (i) $\epsilon = 0.3$

Figure 4.2: The perturbation is $V(q) = \epsilon q_1 q_3^2$ with ϵ ranging from 0.05 to 0.3, the initial energy in the q_2, p_2 system is 0.2 out of a total of 5. The red line is the boundary of the possible points on the $q_3 = 0$ plane. In the latter figures, the phase space is not compact and the points to the left of the verticle line can escape to negative infinity while maintaining a total energy of 5.

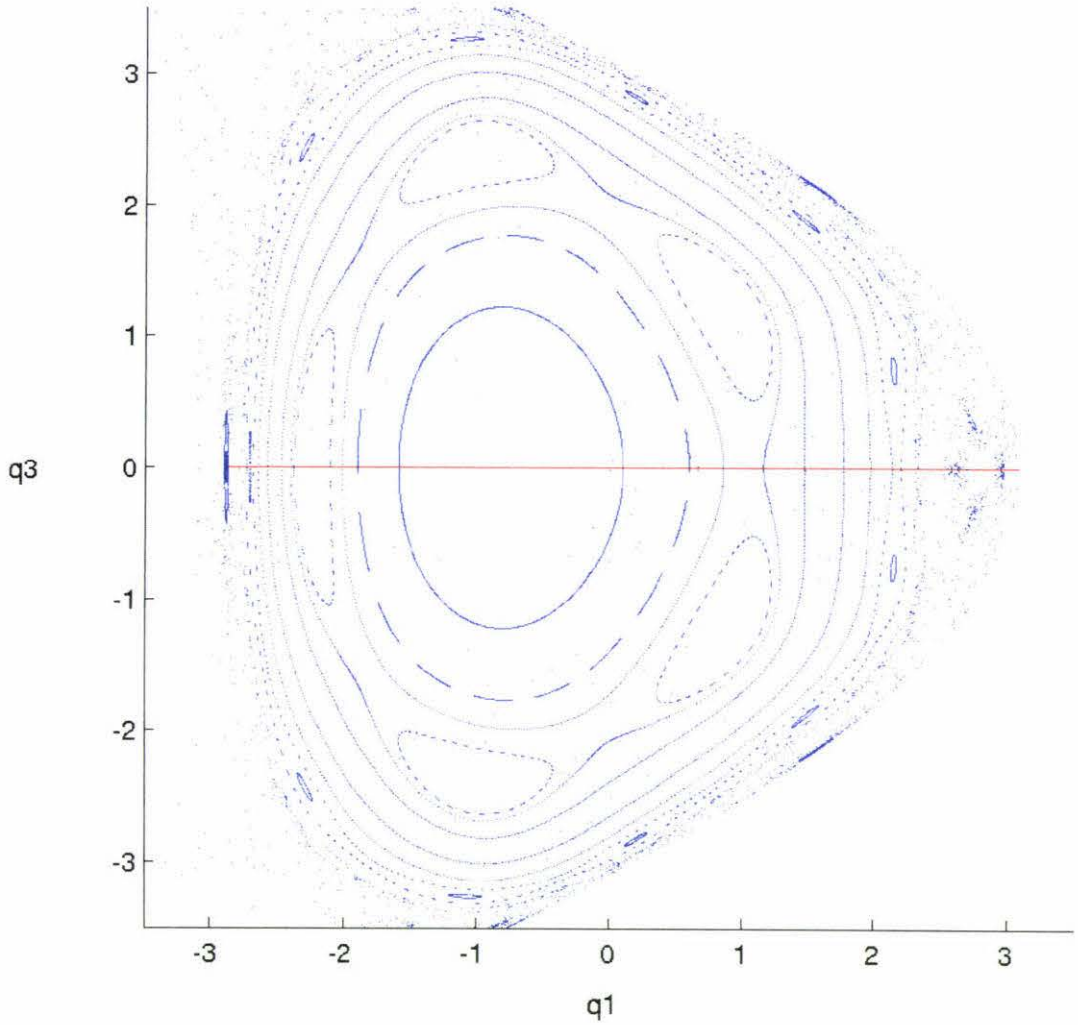
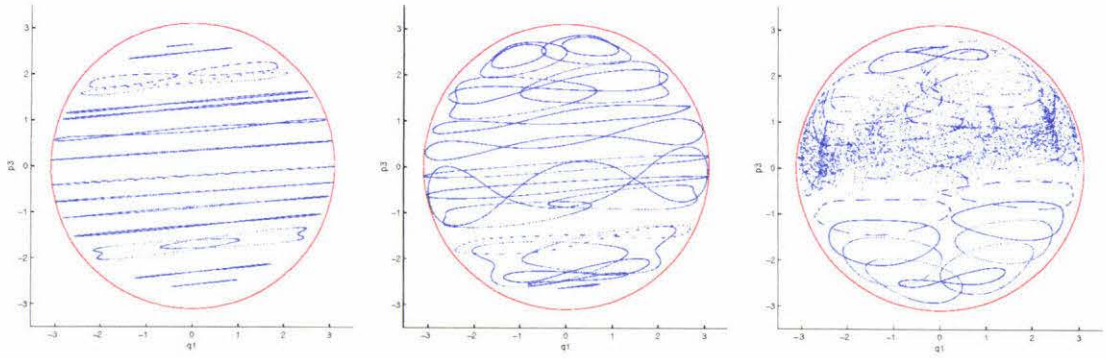


Figure 4.3: All the orbits are symmetric with respect to the 4th reversing symmetry, which prevents any dissipative behaviour. This is the same as figure 4.2(g) but viewed from underneath.

In this system the q_2, p_2 subsystem no longer decouples, allowing energy to be transferred from the p_2 coordinates to the q_1, p_3, q_3 coordinates, keeping $H(q,p) = h_0$. The orbits in the Poincaré section are now free to move in 3 dimensions within the deformed sphere, increasing the chances of dissipative behaviour.

This perturbation breaks the 3rd and 4th symmetries and reversing symmetries but not the 1st or 2nd. For this perturbation, figure 4.4 shows how the sphere distorts as ϵ is increased.



(a) $\epsilon = .001$ (b) $\epsilon = .01$ (c) $\epsilon = .03$

Figure 4.4: Poincaré maps of the $\epsilon q_2 q_3^2$ perturbation as ϵ is increased.

As with the previous perturbation, all the orbits of the Poincaré map appear to be symmetric with respect to some symmetry. This can be best seen in figure 4.5.

The occurrence of the extra symmetry, in each perturbation, was one of the main reasons why, for a long time, I got perplexing results, which I will now discuss.

To begin with I naively did not realise that the system had more reversing symmetries than the anti-symplectic one, $R : p \mapsto -p$, and even once I realised this, it took quite some time before I realised the effect that they were having on the dynamics of the system.

The dynamics of the unperturbed harmonic oscillator system is a rotation about an axis which I calculated in section 3.4.3. But obviously from figure 3.5 it can be seen that the Poincaré map is not symmetric with respect to R. Instead it is symmetric with respect to some other reversing symmetry, which almost all of the orbits cross. So almost all the orbits of this system are symmetric with respect to this reversing symmetry and the ω- and α-limit sets of almost every point must be the same. Thus, there can be no symptoms of dissipation, such as sources, sinks or saddle points with eigenvalues that do not multiply to give 1.

Once I realised the effect that the extra symmetries were having on the dynamics of the system I included the condition that the perturbation must break all the known symmetries of the system to eliminate any effect that the extra symmetries were having on the dynamics. The only reversing symmetry that cannot be broken by a perturbation is the anti-symplectic one as the potential is a function of q_1, q_2 and q_3 only.

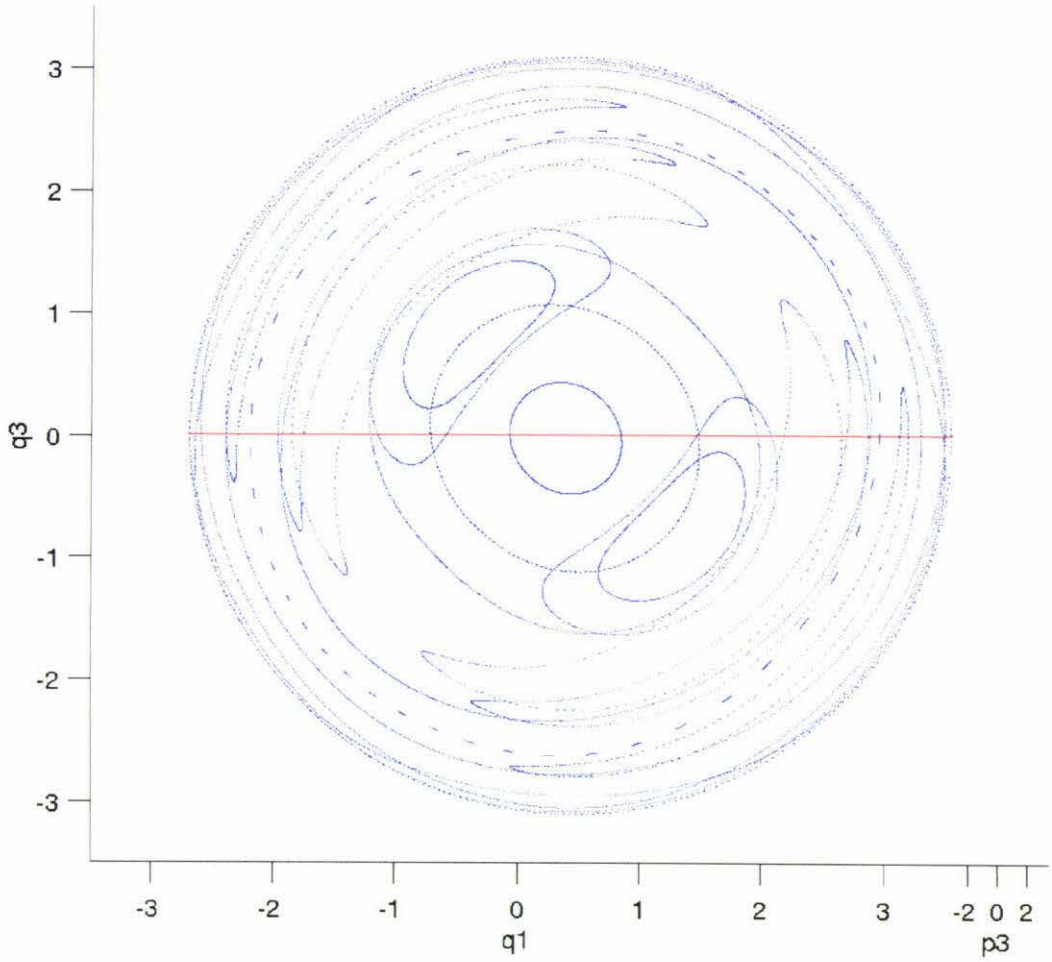


Figure 4.5: A view along the axis of rotation of the unperturbed system showing symmetry of the perturbed system. The figure is the same as figure 4.4(b), with $\epsilon = .01$.

The other main reason that I was getting perplexing results with the harmonic oscillator system is that the Poincaré section was not invariant with respect to the reversing symmetry, $R : (q,p) \mapsto (q,-p)$, of the system. My Poincaré section was

$$\Sigma = \{(q,p) : H(q,p) = h_0, p_1 = 0, q_2 = 0, p_2 > 0\}$$

and I was observing the following:

The orbit of a point x, in the Poincaré map, was the same as the orbit of the point $R(x)$, despite the orbit not crossing the $p_3 = 0$ plane in the Poincaré map.

The cause of the problem is that if the Poincaré section does not contain the fixed set of a symmetry then the orbits that cross that fixed set will appear non-symmetric in the Poincaré map, but will, in fact, be symmetric in the underlying system and not have any dissipative behaviour. This is because of the following:

Let the dimension of the constraint manifold be n, i.e. $\dim M = n$. Then the dimension of the fixed set of R obeys

$$1 \leq \dim \mathrm{Fix}(R) \leq n-1,$$

and the dimension of the Poincaré section is

$$\dim \Sigma = n - 1.$$

Theorem 11 *Let $R : M \mapsto M$ be a reversing symmetry of a reversible system. If $R(\Sigma) = \Sigma$, then $\phi : \Sigma \mapsto \Sigma$ is $R|_\Sigma$-reversible and $\mathrm{Fix}(R|_\Sigma) = \mathrm{Fix}(R) \cap \Sigma$. Then*

$$\dim \mathrm{Fix}(R|_\Sigma) \begin{cases} = \dim \mathrm{Fix}(R) & \text{if } \mathrm{Fix}(R) \subset \Sigma \quad (a) \\ < \dim \mathrm{Fix}(R) & \text{if } \mathrm{Fix}(R) \not\subset \Sigma \quad (b) \end{cases}$$

Proof:

Let $R(\Sigma) = \Sigma$. Since ϕ_t is R-reversible in the reversible system, then $\phi : \Sigma \mapsto \Sigma$ is $R|_\Sigma$-reversible and $\mathrm{Fix}(R|_\Sigma) = \mathrm{Fix}(R) \cap \Sigma$.

- If $\mathrm{Fix}(R) \subset \Sigma$ then $\mathrm{Fix}(R) \cap \Sigma = \mathrm{Fix}(R)$ and hence $\dim \mathrm{Fix}(R|_\Sigma) = \dim \mathrm{Fix}(R)$.
- If $\mathrm{Fix}(R) \not\subset \Sigma$ then $\mathrm{Fix}(R) \cap \Sigma \subset \mathrm{Fix}(R)$ and hence $\dim \mathrm{Fix}(R|_\Sigma) < \dim \mathrm{Fix}(R)$.

On the other hand, if $R(\Sigma) \neq \Sigma$, then there is no obvious candidate for a reversing symmetry of ϕ. The Poincaré section suffers from this problem. The map ϕ inherits reversible-like behaviour from the flow, but it is unclear as to what the reversing symmetry is.

Since it is preferable not to have any hidden symmetries in the Poincaré section, I later changed the Poincaré section to be:

$$\Sigma = \{(q,p) : H(q,p) = h_0, p_2 = 0, \dot{p}_2 > 0, p_1 = -q_2 p_3\},$$

for which $R(\Sigma) = \Sigma$ and $\mathrm{Fix}(R|_\Sigma) = \{(q,p) : (q,p) \in \Sigma, p_3 = 0\}$.

Unfortunately, I did not realise at the time the effect that the Poincaré section and the extra symmetries were having on the dynamics of the Poincaré map and it was decided that there was no simple perturbation that would cause the existence of non-symmetric orbits for this system. So, based on this assumption I changed the potential of the base system from a spherically symmetric potential to being a free particle in the q_3 direction.

4.3 Free particle

In this system there is no q_3 value in the unperturbed potential, thus my system is now:

$$\begin{aligned} V &= \tfrac{1}{2}(q_1^2 + q_2^2) \\ \therefore H &= \tfrac{1}{2}(\|p\|^2 + q_1^2 + q_2^2) \\ \dot{q} &= p \\ \dot{p_1} &= -q_1 + \lambda \\ \dot{p_2} &= -q_2 \\ \dot{p_3} &= \lambda q_2 \end{aligned}$$

with the constraint

$$p_1 + q_2 p_3 = 0$$

and I am treating q_3 as an angle variable. Thus the new phase space is

$$\begin{aligned} q &\in \mathbb{R}^2 \times \mathbb{S}^1 \\ p &\in \mathbb{R}^3 \end{aligned}$$

The unperturbed system can be solved by a method similar to that of section 3.4.3; in fact, we can find the exact Poincaré map analytically.

Let $q_2 = A\cos(t)$ and $p_2 = -A\sin(t)$. Then $p_1 = -q_2 p_3 = -A\cos(t)p_3$ and the system can be written as

$$\begin{aligned} \dot{p}_1 &= -q_1 + \lambda \\ \dot{q}_1 &= -A\cos(t)p_3 \\ \dot{p}_3 &= \lambda A\cos(t) \\ \dot{q}_3 &= p_3. \end{aligned} \tag{4.2}$$

We can calculate λ from the $\dot{p}_1$ equation and the constraint.

$$\begin{aligned} \lambda &= q_1 + \dot{p}_1 \\ &= q_1 - \dot{q}_2 p_3 - q_2 \dot{p}_3 \\ &= q_1 + A\sin(t)p_3 - A^2\cos^2(t)\lambda \\ \therefore \lambda &= \tfrac{q_1 + A\sin(t)p_3}{1 + A^2\cos^2(t)}. \end{aligned}$$

So the $\dot{p}_3$ equation becomes

$$\dot{p}_3 = (1 + A^2\cos^2(t))^{-1} A\cos(t)(q_1 + A\sin(t)p_3)$$

and we can make the substitution $\tilde{p}_3 = (1 + A^2\cos^2(t))^{\frac{1}{2}}$. Our system of equations can now be written as

$$\begin{aligned} \dot{q}_1 &= \alpha\tilde{p}_3 \\ \dot{\tilde{p}}_3 &= -\alpha q_1 \\ \dot{q}_3 &= \beta\tilde{p}_3 \end{aligned} \tag{4.3}$$

where

$$\begin{aligned} \alpha &= -A\cos(t)(1 + A^2\cos^2(t))^{-\frac{1}{2}} \\ \beta &= (1 + A^2\cos^2(t))^{-\frac{1}{2}}. \end{aligned}$$

Now, the $q_1, \tilde{p}_3$ subsystem is a rotation by

$$\Theta = \int_0^t \alpha \mathrm{d}t,$$

so

$$\begin{pmatrix} q_1 \\ \tilde{p}_3 \end{pmatrix} = \begin{pmatrix} \cos(\Theta) & \sin(\Theta) \\ -\sin(\Theta) & \cos(\Theta) \end{pmatrix} \begin{pmatrix} q_1(0) \\ \tilde{p}_3(0) \end{pmatrix}$$

but

$$\int_0^{2\pi} \alpha \mathrm{d}t = 0,$$

so the Poincaré map becomes

$$\begin{aligned} q_1 &\mapsto q_1 \\ p_3 &\mapsto p_3, \end{aligned}$$

while,

$$\begin{aligned} q_3 &\mapsto q_3 + \textstyle\int_0^{2\pi} \beta(t)\tilde{p}_3(t)\mathrm{d}t \\ &= q_3 + \textstyle\int_0^{2\pi} \beta(t)(-\sin(\Theta(t))q_1(0) + \cos(\Theta(t))\tilde{p}_3(0))\mathrm{d}t \\ &= q_3 + \frac{2\pi}{\sqrt{1+A^2}}\tilde{p}_3(0) \\ &= q_3 + 2\pi p_3(0). \end{aligned}$$

So the orbits in the Poincaré map $\phi(q_1, p_3, q_3) = (q_1, p_3, q_3 + 2\pi p_3)$, are invariant circles in the q_3 direction.

The symmetries of this system are:

$$\begin{aligned} 1 \quad & (q_1, q_2, q_3, p_1, p_2, p_3) \mapsto (q_1, q_2, q_3, p_1, p_2, p_3) \qquad (I) \\ 2 \quad & (q_1, q_2, q_3, p_1, p_2, p_3) \mapsto (-q_1, q_2, q_3, -p_1, p_2, -p_3) \\ 3 \quad & (q_1, q_2, q_3, p_1, p_2, p_3) \mapsto (q_1, -q_2, q_3, p_1, -p_2, -p_3) \\ 4 \quad & (q_1, q_2, q_3, p_1, p_2, p_3) \mapsto (-q_1, -q_2, q_3, -p_1, -p_2, p_3) \end{aligned}$$

and the reversing symmetries are:

$$\begin{aligned} 1 \quad & (q_1, q_2, q_3, p_1, p_2, p_3) \mapsto (q_1, q_2, q_3, -p_1, -p_2, -p_3) \qquad (R) \\ 2 \quad & (q_1, q_2, q_3, p_1, p_2, p_3) \mapsto (-q_1, q_2, q_3, p_1, -p_2, p_3) \\ 3 \quad & (q_1, q_2, q_3, p_1, p_2, p_3) \mapsto (q_1, -q_2, q_3, -p_1, p_2, p_3) \\ 4 \quad & (q_1, q_2, q_3, p_1, p_2, p_3) \mapsto (-q_1, -q_2, q_3, p_1, p_2, -p_3). \end{aligned}$$

Any (reversing) symmetry together with $q_3 \mapsto f(q_3)$ is also a (reversing) symmetry.

One class of ω- and α-limit sets is that of a closed loop in phase space. These closed loops appear as fixed or periodic points in the Poincaré map. Near fixed points the dynamics of the system can be linearised as described in sections 2.6-2.8 and the eigenvalues of the linearised coefficient matrix can be computed. If dissipation occurs near these points, then the eigenvalues of the linearised coefficient matrix at those points will multiply to give a number other than 1. For period n points, the linearised coefficient matrix of the nth mapping of the periodic point is taken, i.e. $A = \mathrm{D}(f^n)(x)$.

Since I was looking for dissipative behaviour in the system, searching for the fixed points was the logical place to start. But once again the Poincaré section affected our judgement and I thought that all the fixed points lay within the fixed set of a symmetry and were hence symmetric. So I began searching for period 2 points and computed the eigenvalues of the linearised coefficient matrices to see if there was any dissipation.

To do this I used the code in appendix A.2.4 as the input function to the MATLAB routine *fsolve* which finds zeros of the input function. The routine is called several times with decreasing time-steps as we home-in on the period 2 points. Then I used the code in appendix A.2.5 to compute the eigenvalues of the linearisation at each period 2 point.

For the unperturbed system, the fixed set of R is along $p_3 = 0$ and the period 2 points in the Poincaré map are horizontal lines at $p_3 = \frac{1}{2}(2n-1)$, $n \in \mathbb{Z}$ as shown in figure 4.6. The fixed points occur along horizontal lines at $p_3 = n$, $n \in \mathbb{Z}$. Both the fixed points and the period 2 points have eigenvalues $\lambda_1 = \lambda_2 = \lambda_3 = 1$.

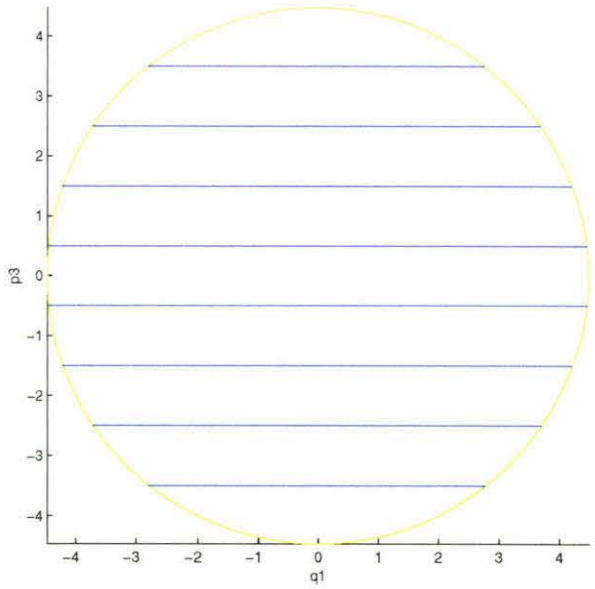


Figure 4.6: Period 2 points with the Poincaré section at $p_2 = 0$, $\dot{p}_2 > 0$. All orbits are either periodic or quasi-periodic, with frequencies 1 and p_3.

4.3.1 Perturbations of the Free Particle

When perturbing this system, I had to take into consideration a few extra conditions. Firstly, I was treating the q_3 variable as an angle so the perturbation had to involve trigonometric functions of q_3 and polynomial functions of q_1 and q_2 only. Secondly, as mentioned in section 4.2.1, the perturbation should break all the known symmetries of the system, with the exception of the unbreakable reversing symmetry R. Thirdly, it should also preserve some of the period 2 points, to facilitate the searching for dissipative behaviour. Preferably, it should make them isolated period 2 points, as continuous curves of period 2 points have an eigenvalue $\lambda = 1$ along the curve and are a result of the system being degenerate in some way.

The size of ϵ that I have used for this system is smaller than that of the harmonic oscillator system, due to having a more complicated perturbation. Generally I used a value between 0.005 and 0.05 for ϵ, with the smaller values not having much effect on the system and the larger values causing too much chaos.

The simplest perturbation that I could find which had only R as a reversing symmetry was:

$$\epsilon(q_1 q_2(\cos(q_3) + \sin(q_3)) + q_1^2 q_2 \cos(q_3))$$

giving the system:

$$\begin{aligned}
\dot{q} &= p \\
\dot{p_1} &= -q_1 + \lambda - \epsilon(q_2(\cos(q_3) + \sin(q_3)) + q_1 q_2 \cos(q_3)) \\
\dot{p_2} &= -q_2 - \epsilon(q_1(\cos(q_3) + \sin(q_3)) + q_1^2 \cos(q_3)) \\
\dot{p_3} &= \lambda q_2 - \epsilon(q_1 q_2(-\sin(q_3) + \cos(q_3)) - q_1^2 q_2 \sin(q_3))
\end{aligned}$$

with the constraint

$$p_1 + q_2 p_3 = 0.$$

The phase portraits obtained typically look like that of figure 4.7. Each orbit of figure 4.7 contains 10000 points, the total energy of the system is $H = 3$, and the perturbation size is $\epsilon = 0.02$. The black orbits are symmetric orbits that are non-rotational in the q_3 direction. The green orbits are also symmetric, but they are rotational in the q_3 direction. The red and cyan orbits appear chaotic and if left long enough, tend to fill out much of the phase space. They all cross the $p_3 = 0$ plane and are symmetric orbits.

The blue orbit, however, is non-symmetric, it is a 2-torus in the Poincaré map and never crosses the $p_3 = 0$ plane. A closeup of this orbit can be seen in figure 4.8. Unfortunately, the ω- and $\alpha(x)$-limit sets appear to be equal, so the orbit is not dissipative.

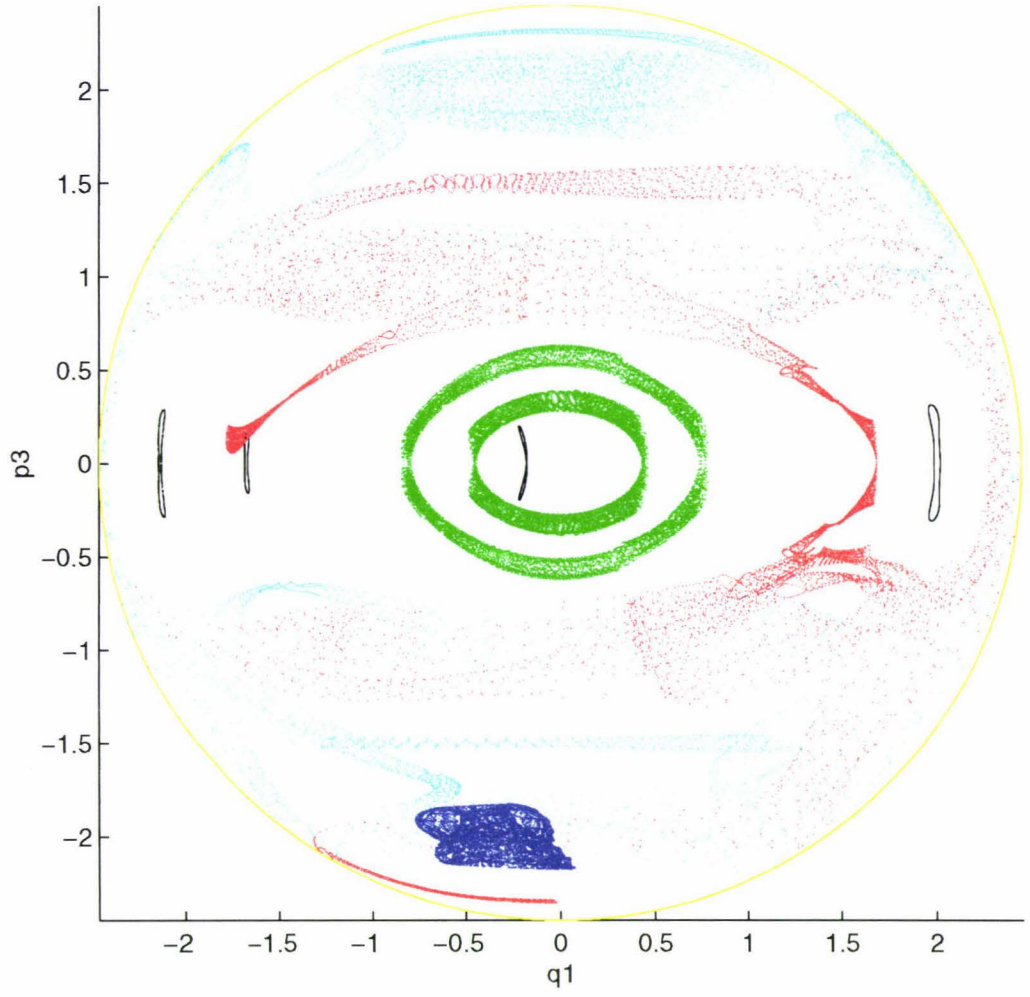
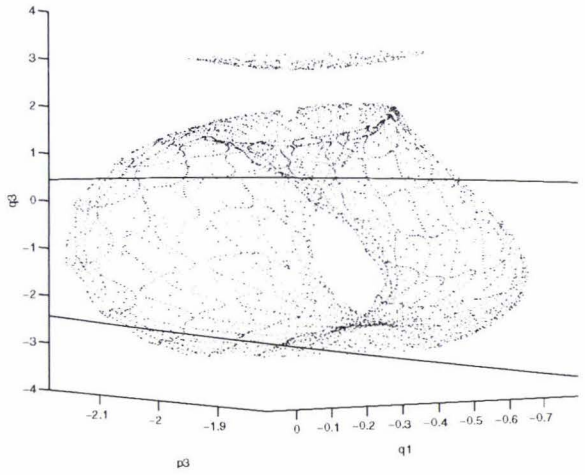
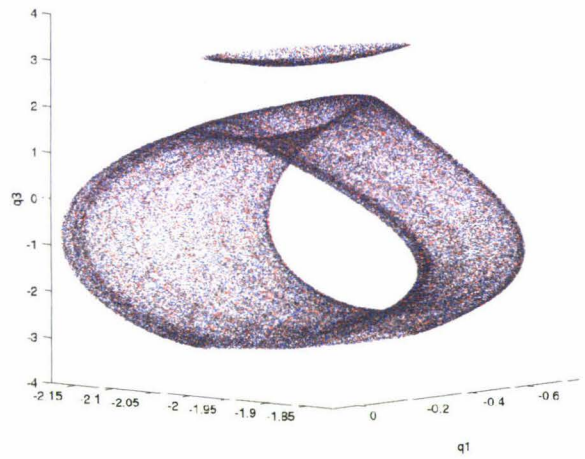


Figure 4.7: A typical graph of the free particle with the perturbation $\epsilon(q_1q_2(\cos(q_3)+\sin(q_3))+q_1^2q_2\cos(q_3))$ showing various orbits in different colours. There are 4 black orbits, 2 green orbits, 1 red orbit, 1 cyan orbit and 1 blue orbit. The total energy is $H = 3$, and the perturbation size is $\epsilon = .02$.



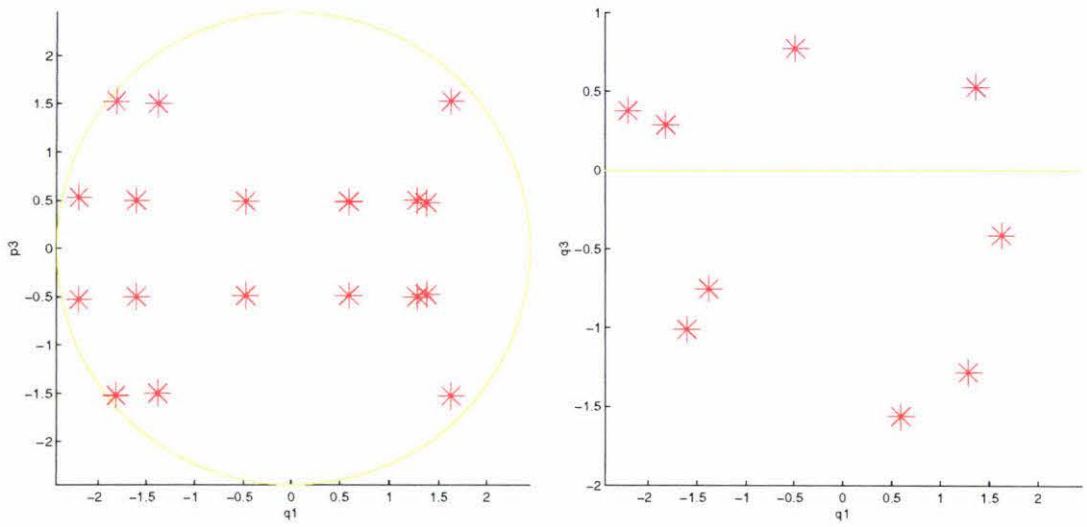
(a) A 2-torus in the Poincaré map with an initial point $(q_1, p_3, q_3) = (0.014094136630869780, -2.159418589032537916, 0)$



(b) The same 2-torus, with 50000 points at a time-step of 0.01 and 50000 points at a time-step of -0.01.

Figure 4.8: A 2-torus in the Poincaré map with an initial point $(q_1, p_3, q_3) = (0.014094136630869780, -2.159418589032537916, 0)$. The total energy is $H = 3$, the perturbation size is $\epsilon = 0.02$. Figure 4.8(a) shows the structure of the 2-torus. Figure 4.8(b) shows how the ω- and α-limit sets occupy the same region of phase space, preventing any dissipation.

The period 2 points that survived perturbation of the free particle system no longer occur in continuous curves but as isolated points in the phase space. The isolated period 2 points that do not lie in the fixed set of the reversing symmetry R had the possibility of being dissipative or expansive. So I computed the eigenvalues of the linearised coefficient matrix of the map at those fixed points to determine their stability.



(a) Front-on view

(b) Side-on view

Figure 4.9: Period 2 points of the perturbed free particle system with perturbation $\epsilon(q_1q_2(\cos(q_3)+\sin(q_3))+q_1^2q_2\cos(q_3))$. The total energy of the system is $H=3$, and the perturbation size $\epsilon=0.02$.

The period 2 points that I have found all have an eigenvalue structure similar to one of the following fixed points

$$(q_1,p_3,q_3)=\begin{pmatrix}1.62616585275446\\-1.52826908140430\\-0.41257126757106\end{pmatrix}\quad\lambda=\begin{pmatrix}0.86718945139869+0.32993449966305i\\0.86718945139869-0.32993449966305i\\1.16159344438038\end{pmatrix}$$

$$(q_1,p_3,q_3)=\begin{pmatrix}1.28505572674583\\-0.50140223444199\\-1.28195494854409\end{pmatrix}\quad\lambda=\begin{pmatrix}0.79284968939701\\1.02985209681468\\1.22471447812289\end{pmatrix}.$$

So these are all spiral saddle points similar to figure 2.4(b) or node saddle points similar to figure 2.4(a). The product of the three eigenvalues for each period 2 point is 1 and thus they do not exhibit dissipative behaviour. Although these have been estimated by finite differences and numerical integration, extensive tests indicate that they are accurate to about 10^{-6}.

The 2-torus and the period 2 points are the only non-symmetric orbits that I have found

in this system, even for small ϵ, despite the unperturbed system having many non-symmetric orbits. This could be a result of the unperturbed system being highly degenerate.

To sum up, the following evidence,

1. the eigenvalues of all periodic points studied have product 1,
2. the only non-symmetric orbit found is an invariant 2-torus and hence has $\alpha = \omega$,

tend to suggest that the system is in fact volume preserving with respect to some unknown volume $m(x)d^n x$. If it exists, this volume is constrained by the dynamics in some way as it is clear that not all nonholonomic systems (e.g. the rattleback) are volume preserving. So far, it is not clear if this is due to the low dimension of the system, the choice of Hamiltonian, or the choice of constraint.

Chapter 5

Results, Conclusions, and Discussions

The main goal of this thesis was to find non-Hamiltonian-like behaviour (specifically, dissipative behaviour) in a class of simple mechanical systems subject to a nonholonomic constraint. Finding dissipative behaviour would strongly support the conjecture that the *only* general features of the dynamics of this class of systems is that they preserve the energy and are reversible.

To integrate these systems on a computer I have written a very fast, simple, reversible, numerical integrator for nonholonomic systems. The class of 6 dimensional nonholonomic systems that I studied, known as contact particle systems in $\mathbb{R}^3$, can be reduced to 3 dimensions via two equations and a Poincaré section. The resulting Poincaré map can be plotted using a program such as MATLAB.

Dissipative orbits can be characterised by differing ω- and α-limit sets. Thus, to search for dissipative behaviour, I plotted the forwards and reverse time orbits of an initial point in different colours and looked for where they occupied different regions of phase space. Since symmetric orbits cannot be dissipative, I perturbed an integrable system having many non-symmetric orbits with the intent of creating a non-integrable system with many non-symmetric orbits.

The initial system was a harmonic oscillator system, of which the Poincaré map is a fixed rotation on the surface of a sphere; the axis of rotation can be calculated with a little help from the Magnus expansion. When perturbed, the sphere deforms and KAM circles appear, these are shortly followed by chaos as the size of the perturbation is increased. The perturbed systems that I considered had extra symmetries and my choice of the Poincaré section did not contain $\mathrm{Fix}(R)$, which was the cause of many perplexing results.

The second system I considered was a free particle system. I changed my choice of the Poincaré section to contain $\mathrm{Fix}(R)$ so as to avoid the confusion that occurred in the previous system. For this system, the Poincaré map can be found analytically and gives invariant circles in the q_3 direction. Perturbations of this system had the requirement that the only reversing

symmetry is the anti-symplectic one, $R : (q,p) \mapsto (q,-p)$. The simplest perturbation that I could find, satisfying the above requirement, was $\epsilon(q_1q_2(\cos(q_3)+\sin(q_3))+q_1^2q_2\cos(q_3))$, but this perturbation makes almost all of the orbits symmetric, even for small ϵ. I suspect that this is due to the degeneracy of the unperturbed system.

Unfortunately, I did not manage to find any evidence of non-Hamiltonian-like behaviour. In the systems that I have studied, all of the periodic points have product 1 and the only non-symmetric orbit found is an invariant 2-torus and hence has $\alpha = \omega$. This lack of dissipation tends to suggest that the systems studied in this thesis preserve some volume element which is not readily apparent. If this volume element exists, then it must be constrained by the dynamics in some way as it is clear that not all nonholonomic systems (e.g. the rattleback) are volume preserving.

Considering the various perturbations that I have studied, it is unlikely that all of the systems in this thesis are special, volume preserving cases of contact particle systems. But, at this point in time it is not clear as to the cause of the volume preservation. Possibly, the occurrence of volume preservation could be due to the low dimension of system, the choice of Hamiltonian, the choice of constraint, or a combination of the above. In any case, further research into mechanical systems subject to nonholonomic constraints is required to establish the validity of the above conjecture.

Appendix A

Program Code

A.1 Mex Files

A.1.1 Integrator Code

```
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <math.h>
#include "mex.h"
/*
Calling syntax: [q1, q2, q3, p1, p2, p3, Rotations, StepsTaken] = ...
                Free_eps_p2(in, NumLoops, H, dt, eps, SuppressOutput)
eps(ilon) is the size of the perturbation

*/

/* Input Arguments */

#define in_IN  prhs[0]
#define NumLoops_IN  prhs[1]
#define Energy_IN  prhs[2]
#define TimeStep_IN  prhs[3]
#define eps_IN prhs[4]
#define SuppressOutput_IN prhs[5]

/* Output Arguments */

#define sectionq1_OUT plhs[0]
#define sectionq2_OUT plhs[1]
#define sectionq3_OUT plhs[2]
#define sectionp1_OUT plhs[3]
#define sectionp2_OUT plhs[4]
#define sectionp3_OUT plhs[5]
```

```
#define Rotations_OUT plhs[6]
#define StepsTaken_OUT  plhs[7]


void NHIntegrator(double *, double *, double *, double *, double *, double *,...
                  double *, double*, double, double, long, double, int,...
                  double *, double *);
void Step(double *, double *, double, double);
void nhpsmain(double *, double *, double *, double *, double *, double *,...
              double *, long , double , double, double, int, double *, double *);

/* Entry point of "mex" function */

void mexFunction(int nlhs, mxArray *plhs[], int nrhs, const mxArray *prhs[])
{
  double *sectionq1, *sectionq2, *sectionq3;
  double *sectionp1, *sectionp2, *sectionp3, *in;
  double *Energy, *H, *dt, *eps, *Rotations, *StepsTaken;
  double *NumLoops, *SuppressOutput;

/* Check for proper number of arguments */

  if (nrhs != 6)
    {
      mexErrMsgTxt("Free_eps_p2 requires six input arguments.");
    }
  else if (nlhs > 8)
    {
      mexErrMsgTxt("Free_eps_p2 has only eight output arguments.");
    }

/* Check the dimensions of in.  in must be 3 X 1. */

  if (!mxIsNumeric(in_IN) || mxIsComplex(in_IN) ||
      mxIsSparse(in_IN)  || !mxIsDouble(in_IN) ||
      (mxGetM(in_IN) != 3) || (mxGetN(in_IN) != 1))
    {
      mexErrMsgTxt("nhps requires that 'in' be a 3 x 1 real vector.");
    }

/* Create matrices for the return arguments */

  NumLoops = mxGetPr(NumLoops_IN);
  sectionq1_OUT = mxCreateDoubleMatrix(*NumLoops, 1, mxREAL);
  sectionq2_OUT = mxCreateDoubleMatrix(*NumLoops, 1, mxREAL);
  sectionq3_OUT = mxCreateDoubleMatrix(*NumLoops, 1, mxREAL);
  sectionp1_OUT = mxCreateDoubleMatrix(*NumLoops, 1, mxREAL);
  sectionp2_OUT = mxCreateDoubleMatrix(*NumLoops, 1, mxREAL);
  sectionp3_OUT = mxCreateDoubleMatrix(*NumLoops, 1, mxREAL);
```

```
  Rotations_OUT = mxCreateDoubleMatrix(1,1,mxREAL);
  StepsTaken_OUT = mxCreateDoubleMatrix(1,1,mxREAL);

/* Assign pointers to the various parameters */

  SuppressOutput = mxGetPr(SuppressOutput_IN);
  sectionq1 = mxGetPr(sectionq1_OUT);
  sectionq2 = mxGetPr(sectionq2_OUT);
  sectionq3 = mxGetPr(sectionq3_OUT);
  sectionp1 = mxGetPr(sectionp1_OUT);
  sectionp2 = mxGetPr(sectionp2_OUT);
  sectionp3 = mxGetPr(sectionp3_OUT);
  in = mxGetPr(in_IN);
  H = mxGetPr(Energy_IN);
  dt = mxGetPr(TimeStep_IN);
  eps = mxGetPr(eps_IN);
  Rotations = mxGetPr(Rotations_OUT);
  StepsTaken = mxGetPr(StepsTaken_OUT);

/* Do the actual computations in a subroutine */

  nhpsmain(sectionq1, sectionq2, sectionq3, sectionp1, sectionp2, sectionp3,...
           in, (int) *NumLoops, *H, *dt, *eps, (int) *SuppressOutput,...
           Rotations, StepsTaken);
  return;
}

void nhpsmain(double *sectionq1, double *sectionq2, double *sectionq3,...
              double *sectionp1, double *sectionp2, double *sectionp3,...
              double *in, long NumLoops, double H, double dt, double eps,...
              int SuppressOutput, double *Rotations, double *StepsTaken)
{
  double p[3],q[3],a,b,c;
  double E;
  *Rotations=0;
  *StepsTaken=0;

  q[0]=in[0];
  q[2]=in[2];
  p[2]=in[1];
  p[1]=0;
/* H=.5(|p|^2+q1^2+q2^2)+eps*(q1*q2*(cos(q3)+sin(q3))+q1^2*q2*cos(q3))
   p1=-q2*p3

   (1+p3^2)q2^2+(2eps*(q1*(cos(q3)+sin(q3))+q1^2*cos(q3)))q2+(p2^2+p3^2+q1^2-2H)=0
*/
  a=1+p[2]*p[2];
  b=2*eps*(q[0]*(cos(q[2])+sin(q[2]))+q[0]*q[0]*cos(q[2]));
  c=-2*H+p[1]*p[1]+p[2]*p[2]+q[0]*q[0];
```

```
  if ((-b+sqrt(b*b-4*a*c))/(2*a)<0)
    {
      printf("q2 has 2 possible values!\n");
      return;
    }
  q[1]=(-b-sqrt(b*b-4*a*c))/(2*a);
  p[0]=-q[1]*p[2];

  if(!SuppressOutput)
    printf("H is %.18f\nq1 is %.18f\nq2 is %.18f\nq3 is %.18f\np1 is %.18f\n"...
           "p2 is %.18f\np3 is %.18f\n",H,q[0],q[1],q[2],p[0],p[1],p[2]);

  E=.5*(p[0]*p[0]+p[1]*p[1]+p[2]*p[2]+q[0]*q[0]+q[1]*q[1])+...
        eps*(q[0]*q[1]*(cos(q[2])+sin(q[2]))+q[0]*q[0]*q[1]*cos(q[2]));
  if(!SuppressOutput)
    printf("Initial energy = %.18f\n",E);

  if (q[1]<0)
    NHIntegrator(q, p, sectionq1, sectionq2, sectionq3, sectionp1, sectionp2,...
                 sectionp3, dt, H, NumLoops, eps, SuppressOutput, Rotations,...
                 StepsTaken);
  if(SuppressOutput!=2)
    printf("Number of rotations in the q3 direction was %.0f\n",*Rotations);
}

void NHIntegrator(double *q, double *p, double *sectionq1, double *sectionq2,...
                  double *sectionq3, double *sectionp1, double *sectionp2,...
                  double *sectionp3, double dt, double H, long NumLoops,...
                  double eps, int SuppressOutput, double *Rotations,...
                  double *StepsTaken)
{
  int i=0;
  int flag=0,flag2=0;
  double o,E,Edrift=0.0,E2,Edrift2=0.0;
  double so[6];
  if(!SuppressOutput)
    printf("Doing section\n");
  while((i<NumLoops))
    {
/* Retain old values to find Poincare section */
      so[0]=q[0];
      so[1]=q[1];
      so[2]=q[2];
      so[3]=p[0];
      so[4]=p[1];
      so[5]=p[2];
      Step(q,p,dt,eps);
      (*StepsTaken)++;
/* Count the number of rotations in the q3 direction */
```

```
      if(q[2]>M_PI)
        {
          if(flag==0)flag=1;
          q[2]=fmod(q[2]+M_PI,2*M_PI)-M_PI;
          *Rotations+=flag;
        }
      else if(q[2]<=-M_PI)
        {
          if(flag==0)flag=-1;
          q[2]=fmod(q[2]-M_PI,2*M_PI)+M_PI;
          *Rotations-=flag;
        }

/* Test condition of the Poincare section */
      if (((dt*so[4]<=0)&&(dt*p[1]>0)))/*||((so[4]>=0)&&(p[1]<0)))*/
        {
          sectionq1[i]=q[0]-(q[0]-so[0])*p[1]/(p[1]-so[4]);
          sectionq2[i]=q[1]-(q[1]-so[1])*p[1]/(p[1]-so[4]);
          sectionq3[i]=q[2]-(q[2]-so[2])*p[1]/(p[1]-so[4]);
          sectionp1[i]=p[0]-(p[0]-so[3])*p[1]/(p[1]-so[4]);
          sectionp2[i]=p[1]-(p[1]-so[4])*p[1]/(p[1]-so[4]);
          sectionp3[i]=p[2]-(p[2]-so[5])*p[1]/(p[1]-so[4]);
          E=.5*(sectionp1[i]*sectionp1[i]+sectionp2[i]*sectionp2[i]+sectionp3[i]*...
               sectionp3[i]+sectionq1[i]*sectionq1[i]+sectionq2[i]*sectionq2[i])+...
            eps*(sectionq1[i]*sectionq2[i]*(cos(sectionq3[i])+sin(sectionq3[i]))+...
                sectionq1[i]*sectionq1[i]*sectionq2[i]*cos(sectionq3[i]));

          E2=.5*(p[0]*p[0]+p[1]*p[1]+p[2]*p[2]+q[0]*q[0]+q[1]*q[1])+...
                 eps*(q[0]*q[1]*(cos(q[2])+sin(q[2]))+q[0]*q[0]*q[1]*cos(q[2]));
          Edrift+=sqrt((E-H)*(E-H));
          Edrift2+=sqrt((E2-H)*(E2-H));
          i++;
          if(mxIsNaN(E))break;
        }
    }
  if(!SuppressOutput)
    {
      printf("Section energy drift = %.18f\n",Edrift/NumLoops);
      printf("Last point energy drift = %.18f\n",Edrift2/NumLoops);
      printf("Final Energy = %.18f\n",E);
      printf("Number of Steps = %.0f\n",*StepsTaken);
    }
}

void Step(double *q, double *p, double dt, double eps)
{
  double x,y,z,Vx,Vy,Vz,lambda,yn;
  double pt[3],qt[3];
  qt[0]=q[0]+.5*dt*p[0];
```

```
  qt[1]=q[1]+.5*dt*p[1];
  qt[2]=q[2]+.5*dt*p[2];
  x=qt[0];y=qt[1];z=qt[2];

/* H=.5(|p|^2+q1^2+q2^2)+eps*(q1*q2*(cos(q3)+sin(q3))+q1^2*q2*cos(q3))
*/
  Vx=x+eps*(y*(cos(z)+sin(z))+2*x*y*cos(z));
  Vy=y+eps*(x*(cos(z)+sin(z))+x*x*cos(z));
  Vz=eps*(x*y*(-sin(z)+cos(z))-x*x*y*sin(z));
  pt[0]=p[0]-dt*Vx;
  pt[1]=p[1]-dt*Vy;
  pt[2]=p[2]-dt*Vz;
  yn=y+.5*dt*pt[1];
  lambda=-(pt[0]+yn*pt[2])/(1+yn*y);
  p[0]=pt[0]+lambda;
  p[1]=pt[1];
  p[2]=pt[2]+lambda*y;
  q[0]=qt[0]+.5*dt*p[0];
  q[1]=qt[1]+.5*dt*p[1];
  q[2]=qt[2]+.5*dt*p[2];
}
```

A.1.2 Density Map

```
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <math.h>
#include "mex.h"
/*
  Calling syntax: [density] = density(q, NumLoops, H, dt, eps, P2Symm, gridsize)
  eps(ilon) is the size of the perturbation
*/

/* Input Arguments */

#define q_IN  prhs[0]
#define NumLoops_IN  prhs[1]
#define Energy_IN  prhs[2]
#define TimeStep_IN  prhs[3]
#define eps_IN prhs[4]
#define P2Symm_IN prhs[5]
#define gridsize_IN  prhs[6]

/* Output Arguments */

#define density_OUT plhs[0]
```

```
void NHIntegrator(double *, double *, double *, double , long , double , double);
void Step(double *, double *, double, double);
void nhpsmain(double *, double *, long , double , double , double , double ,...
              double);

void mexFunction(int nlhs, mxArray *plhs[], int nrhs, const mxArray *prhs[])
{
  double *in;
  double *Energy, *H, *dt, *eps;
  double *NumLoops, *P2Symm;
  double *density;
  double *gridsize;

/* Check for proper number of arguments */

  if (nrhs != 7)
    {
      mexErrMsgTxt("nhps requires seven input arguments.");
    }
  else if (nlhs > 1)
    {
      mexErrMsgTxt("nhps has one output argument.");
    }

/* Check the dimensions of q.  q must be 3 X 1. */

  if (!mxIsNumeric(q_IN) || mxIsComplex(q_IN) ||
      mxIsSparse(q_IN)  || !mxIsDouble(q_IN) ||
      (mxGetM(q_IN) != 3) || (mxGetN(q_IN) != 1))
    {
      mexErrMsgTxt("nhps requires that 'in' be a 3 x 1 real vector.");
    }

/* Create matrices for the return arguments */

  gridsize = mxGetPr(gridsize_IN);
  NumLoops = mxGetPr(NumLoops_IN);
  density_OUT = mxCreateDoubleMatrix(((int)((int)((int)(*gridsize))*(*gridsize))*...
                                     (*gridsize)), 1, mxREAL);

/* Assign pointers to the various parameters */

  P2Symm = mxGetPr(P2Symm_IN);
  density = mxGetPr(density_OUT);
  in = mxGetPr(q_IN);
  H = mxGetPr(Energy_IN);
  dt = mxGetPr(TimeStep_IN);
  eps = mxGetPr(eps_IN);
```

```
/* Do the actual computations in a subroutine */

  nhpsmain(density, in, (long) *NumLoops, *H, *dt, *eps, *P2Symm, *gridsize);
  return;
}

void nhpsmain(double *density, double *in, long NumLoops, double H, double dt,...
              double eps, double P2Symm, double gridsize)
{
  double p[3],q[3];
  double E;

  q[0]=in[0];
  p[2]=in[1];
  q[2]=in[2];
  p[0]=0;
  q[1]=0;

  printf("H is %.16f\nq1 is %.16f\np3 is %.16f\nq3 is %.16f\n",H,q[0],p[2],q[2]);
  p[1]=P2Symm*sqrt(2*H-q[0]*q[0]-q[2]*q[2]-p[2]*p[2]-2*eps*q[1]*q[2]*q[2]);
/* H=.5(|p|^2+|q|^2)+eps*q2*q3^2
*/
  printf("p2=%.16f\nA=%.16f\n",p[1], sqrt(p[1]*p[1]+q[1]*q[1]));
  E=.5*(p[0]*p[0]+p[1]*p[1]+p[2]*p[2]+q[0]*q[0]+q[1]*q[1]+q[2]*q[2])+...
    eps*q[1]*q[2]*q[2];
  printf("Initial energy = %.16f\n",E);

  if (2*H-q[0]*q[0]-p[2]*p[2]-q[2]*q[2]-2*eps*q[1]*q[2]*q[2]>=0)
    NHIntegrator(q, p, density, dt, NumLoops, eps, gridsize);
  printf("Done :-)\n");
}

void NHIntegrator(double *q, double *p, double *density, double dt,...
                  long NumLoops, double eps, double gridsize)
{
  int i=0;
  long j,jx,jy,jz;
  double o,E,Edrift=0.0;
  double so[5],section[5];
  printf("Doing section\n");
  while((i<NumLoops))
    {
      o=q[1];
      so[0]=q[0];
      so[1]=p[2];
      so[2]=q[2];
      so[3]=p[0];
      so[4]=p[1];
```

```
      Step(q,p,dt,eps);

/* Test condition of the Poincare section */
      if ((dt*o<=0)&&(dt*q[1]>0))
        {
          section[0]=q[0]-(q[0]-so[0])*q[1]/(q[1]-o);
          section[1]=p[2]-(p[2]-so[1])*q[1]/(q[1]-o);
          section[2]=q[2]-(q[2]-so[2])*q[1]/(q[1]-o);
          section[3]=p[0]-(p[0]-so[3])*q[1]/(q[1]-o);
          section[4]=p[1]-(p[1]-so[4])*q[1]/(q[1]-o);
/* Find the index into the density matrix */
          jx=(long)((section[0]+4)*gridsize/8);
          jy=(long)((section[1]+4)*gridsize/8);
          jz=(long)((section[2]+4)*gridsize/8);
          j=jx+(long)(jy*gridsize)+(long)(jz*gridsize*gridsize);
          if((j<0)||(j>=(int)((int)((int)(gridsize)*gridsize)*gridsize)))
            {
              printf("illegal position\n");
              break;
            }
/* Increase the density */
          density[j]++;
          E=.5*(p[0]*p[0]+p[1]*p[1]+p[2]*p[2]+q[0]*q[0]+q[1]*q[1]+q[2]*q[2])+...
                eps*q[1]*q[2]*q[2];
          Edrift+=sqrt((E-5)*(E-5));
          if(mxIsNaN(E))break;
          i++;
          if(i%((int)(NumLoops/100))==0)
            printf("%d%% ",(int)(100*i/NumLoops));
          if(i%((int)(NumLoops/10))==0)
            printf("\n");
        }
    }
  printf("Energy drift = %.16f\n",Edrift/NumLoops);
  printf("Final Energy = %.16f\n", E);
}

void Step(double *q, double *p, double dt, double eps)
{
  double x,y,z,Vx,Vy,Vz,lambda,yn;
  double pt[3],qt[3];
  qt[0]=q[0]+.5*dt*p[0];
  qt[1]=q[1]+.5*dt*p[1];
  qt[2]=q[2]+.5*dt*p[2];
  x=qt[0];y=qt[1];z=qt[2];
  Vx=x;Vy=y+eps*z*z;Vz=z+eps*2*y*z;
  pt[0]=p[0]-dt*Vx;
  pt[1]=p[1]-dt*Vy;
  pt[2]=p[2]-dt*Vz;
```

```
  yn=y+.5*dt*pt[1];
  lambda=-(pt[0]+yn*pt[2])/(1+yn*y);
  p[0]=pt[0]+lambda;
  p[1]=pt[1];
  p[2]=pt[2]+lambda*y;
  q[0]=qt[0]+.5*dt*p[0];
  q[1]=qt[1]+.5*dt*p[1];
  q[2]=qt[2]+.5*dt*p[2];
}
```

A.2 M files

A.2.1 Magnus Expansion

```
function [Theta] = magnus(B, Orderk, Orderm, Order)
% [Theta] = magnus(B, Orderk, Orderm, Order)
% Finds the Magnus expansion of B to order 'Order' in k
% B must be a matrix of the symbolic variables Xi, A
% The returned matrix, Theta,  is a matrix of A to
% order 'Order'.

syms t;
syms Xi;
syms A;
[sm sn]=size(B);
Theta = zeros(sm,sn);
ThetaOld = Theta;
OrderFix = sprintf('O(A^%d)',Order+1);

Bn=[1 -1/2 1/6 0 -1/30 0 1/42 0 -1/30 0 5/66 0 -691/2730 0 7/6];
% Bn is the first 15 Bernoulli numbers

for m=2:Orderm+1
  m-1
  Theta = zeros(sm,sn);
  for k=0:Orderk
% ignore Bernoulli numbers which are zero
     if Bn(k+1)~=0
% 'pack'
       pack
       temp = AdOld(ThetaOld,B,0);
% try to reduce the terms via a taylor expansion
       try temp2=taylor(temp(1,2)+OrderFix,A,Order+1); temp(1,2)=temp2; end
       try temp2=taylor(temp(1,3)+OrderFix,A,Order+1); temp(1,3)=temp2; end
       try temp2=taylor(temp(2,3)+OrderFix,A,Order+1); temp(2,3)=temp2; end
% OrderFix is added so that the taylor expansion doesn't throw away useful terms
       temp(2,1) = -temp(1,2);
```

```
      temp(3,1) = -temp(1,3);
      temp(3,2) = -temp(2,3);
      temp(1,1) = 0;
      temp(2,2) = 0;
      temp(3,3) = 0;
% perform the adjoint to order k
     for i=1:k
       temp = AdOld(ThetaOld,temp,1);
       try temp2=taylor(temp(1,2)+OrderFix,A,Order+1); temp(1,2)=temp2; end
       try temp2=taylor(temp(1,3)+OrderFix,A,Order+1); temp(1,3)=temp2; end
       try temp2=taylor(temp(2,3)+OrderFix,A,Order+1); temp(2,3)=temp2; end
       temp(2,1) = -temp(1,2);
       temp(3,1) = -temp(1,3);
       temp(3,2) = -temp(2,3);
       temp(1,1) = 0;
       temp(2,2) = 0;
       temp(3,3) = 0;
     end

% integrate the terms and try to reduce them via a taylor expansion
     temp(1,2) = int(temp(1,2),Xi,0,t);
     temp(1,3) = int(temp(1,3),Xi,0,t);
     temp(2,3) = int(temp(2,3),Xi,0,t);
     try temp2=taylor(temp(1,2)+OrderFix,A,Order+1); temp(1,2)=temp2; end
     try temp2=taylor(temp(1,3)+OrderFix,A,Order+1); temp(1,3)=temp2; end
     try temp2=taylor(temp(2,3)+OrderFix,A,Order+1); temp(2,3)=temp2; end
     temp(2,1) = -temp(1,2);
     temp(3,1) = -temp(1,3);
     temp(3,2) = -temp(2,3);
     temp(1,1) = 0;
     temp(2,2) = 0;
     temp(3,3) = 0;
% add in the next highest term with the corresponding Bernoulli coefficient
     if Bn(k+1)./factorial(k)==1
       Theta = Theta + temp;
     else
       Theta = Theta + temp.*Bn(k+1)./factorial(k);
     end
   end
 end
% output the time 2*pi value of Theta
 try
   out=subs(Theta,t,2*pi)-subs(Theta,t,0)
 end
% output the time t value of Theta
 ThetaOld = subs(Theta,t,Xi);
end
```

A.2.2 Adjoint

```
function [Adjoint]=Ad(Theta,A,k)
% Finds the Lie bracket of order k
% i.e. [Theta,[Theta,[Theta,[Theta,[Theta,A]]]]]

 Adjoint=A;
 for i=1:k
   Adjoint = Theta*Adjoint - Adjoint*Theta;
 end
```

A.2.3 Main Controlling Script

```
function fp=Free_near_p2_new()
% Free_near_p2_new() plots the first 1000 points from where you left click, then
% uses fsolve to find the nearest period 2 fixed point. If a fixed point is found
% the eigenvalues of that fixed point will be calculated.
% Middle clicking uses the last value calculated of the fixed point to
% call fsolve several times with decreasing time steps to improve accuracy.
% (This is often needed to find the fixed point in the first place)
% Right clicking exits the program.

global StepsTaken;
H=10;
eps=0.00;

colour=['b.r.g.k.m.'];
figure(2);
%setupgraph(H);
format compact
format long
options=optimset('tolfun',1e-16);%,'tolx',1e-16);
in=[0;0;0];
[in(1) in(2) button]=ginput(1);
in(3)=0;
i=-1;
while button~=3
  StepsTaken=0;
  i=i+2;
  if i>10;i=1;end
  if button==1
/* Integrate the system */
    [q1,q2,q3,p1,p2,p3,Rotations,StepsTaken]=Free_eps_p2_new(in,100000,H,.01,eps,0);
    h=plot3(q1,p3,q3,colour(i:i+1));
    set(h,'MarkerSize',1);drawnow;
/* Preliminary search for periodic point */
    [fp,dist,flag]=fsolve('distof_p2_new',[q1(1);p3(1);q3(1)],options,H,.01,eps);
    dist
    [q1,q2,q3,p1,p2,p3]=Free_eps_p2_new(in,2,H,.01,eps,2);
```

```
    dist2=norm([q1(2),p3(2),q3(2),q2(2)]-[q1(1),p3(1),q3(1),q2(1)])
    fp
    if (flag==1)&(norm(dist)~=0)&(dist2>1e-6)
      h=plot3(fp(1),fp(2),fp(3),'r*');set(h,'MarkerSize',20);
      drawnow;
      [Dphi,V,D]=eigenvals_new(fp,H,eps);
      D,D(1)*D(2)*D(3)
      fp
    end
    StepsTaken
    [in(1) in(2) button]=ginput(1);
    in(3)=0;
  else
    if button==2
/* Search for periodic point at a time step of 0.01 */
      flag=0;dist=1;in=[0;0;0];dt=.01;
      while (flag==0)
        disp('dt=.01')
        in=fp;
        [fp,dist,flag]=fsolve('distof_p2_new',in,options,H,dt,eps)
      end
/* Decrease time step to 0.001 */
      if flag~=1;flag=0;in=[0;0;0];dt=.001;
        while (flag==0)
          disp('dt=.001')
          in=fp;
          [fp,dist,flag]=fsolve('distof_p2_new',in,options,H,dt,eps)
        end
      end
/* Decrease time step to 0.0001 */
      if flag~=1;flag=0;in=[0;0;0];dt=.0001;
        while (flag==0)
          disp('dt=.0001')
          in=fp;
          [fp,dist,flag]=fsolve('distof_p2_new',in,options,H,dt,eps)
        end
      end
/* Decrease time step to 0.00001 (Rarely needed) */
      if flag~=1;flag=0;in=[0;0;0];dt=.00001;
        while (flag==0);
          disp('dt=.00001')
          in=fp;
          [fp,dist,flag]=fsolve('distof_p2_new',in,options,H,dt,eps)
        end
      end
/* Make sure its period 2, not period 1 */
      [q1,q2,q3,p1,p2,p3]=Free_eps_p2_new(in,2,H,dt,eps,2);
      dist2=norm([q1(2),p3(2),q3(2),q2(2)]-[q1(1),p3(1),q3(1),q2(1)])
      if (flag==1)&(norm(dist)~=0)&(dist2>1e-6)
```

```
        h=plot3(fp(1),fp(2),fp(3),'r*');set(h,'MarkerSize',20);drawnow;
        [Dphi,V,D]=eigenvals_new(fp,H,eps);
        D,D(1)*D(2)*D(3)
        fp
      end
      StepsTaken
      [in(1) in(2) button]=ginput(1);
      in(3)=0;
    end
  end
end
```

A.2.4 Distance

```
function [out]=distof_p2_new(in,H,dt,eps)
format long;
global StepsTaken;
[q1,q2,q3,p1,p2,p3,Rotations,StepsTakenThisTime]=Free_eps_p2_new(in,3,H,dt,eps,2);
out=[q1(3),p3(3),q3(3),q2(3)]-[q1(1),p3(1),q3(1),q2(1)];
StepsTaken=StepsTaken+StepsTakenThisTime;
```

A.2.5 Eigenvalues

```
function [Dphi,V,D]=eigenvals_new(in,H,eps)
% Finds the eigenvalues of a period 2 fixed point

format long;
Dphi=zeros(3);
col=['b*' 'r*' 'g*' 'k*' 'y*' 'c*' 'm*'];
[q1,q2,q3,p1,p2,p3]=Free_eps_p2_new(in,3,H,.0001,eps,2);
d=1e-6;
for k=1:3
if k==1;dx=d*[1;0;0];end
if k==2;dx=d*[0;1;0];end
if k==3;dx=d*[0;0;1];end
for j=[1 -1]
[q1,q2,q3,p1,p2,p3]=Free_eps_p2(in+j*dx,3,H,.0001,eps,2);
Dphi(k,:)=Dphi(k,:)+j*[q1(3),p3(3),q3(3)];
end
end
Dphi=Dphi/2/d;
[V,D]=eig(Dphi);
D=[D(1,1);D(2,2);D(3,3)];
```

Bibliography

[1] V.I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer-Verlag, New York, 2nd edition, 1989.

[2] V.I. Arnold, V.V. Kozlov, and A.I. Neishtadt. Mathematical aspects of classical and celestial mechanics. In *Dynamical System III*. Springer-Verlag, Berlin, Heidelberg, 1988.

[3] V.I. Arnold and M.B. Sevryuk. Oscilations and bifurcations in reversible systems. In *Nonlinear Phenomena in Plasma Physics and Hydrodynamics*, pages 31–64. Mir, Moscow, 1986.

[4] D.K. Arrowsmith and C.M. Place. *Dynamical Systems: Differential equations, maps and chaotic behaviour*. Chapman & Hall Mathematics, 1992.

[5] G.D. Birkhoff. *The restricted problem of three bodies*, chapter 39, pages 265–334. Rend. Circ. Mat. Palermo, 1915.

[6] A.M. Bloch. Nonholonomic mechanics and control. Preprint, October 2000.

[7] A.M. Bloch, P.S. Krishnaprasad, J.E. Marsden, and R.M. Murray. Nonholonomic mechanical systems with symmetry. *Archive for rotational mechanics and analysis*, 136(1):21–99, 1996.

[8] L. Boltzmann. Über irreversible strahlungsvorgänge I. *Berl. Ber.*, pages 660–662, 1897a.

[9] L. Boltzmann. Über irreversible strahlungsvorgänge II. *Berl. Ber.*, pages 1016–1018, 1897b.

[10] L. Boltzmann. Über vermeintlich irreversible strahlungsvorgänge II. *Berl. Ber. S*, pages 182–187, 1898.

[11] H. Bondi. The rigid body dynamics of unidirectional spin. *Proc. R. Soc. Lond. A*, 405:265–274, 1986.

[12] F. Cantrijn, M. de Leon, and D.M. de Diego. On almost-poisson structures in nonholonomic mechanics. *Nonlinearity*, 12(3):721–737, 1999.

[13] F. Cantrijn, M. de Leon, J.C. Marrero, and D.M. de Diego. Reduction of nonholonomic mechanical systems with symmetries. *Reports on mathematical physics*, 42(1-2):25–45, 1998.

[14] F. Cardin and M. Favretti. On nonholonomic and vakonomic dynamics of mechanical systems with nonintegrable constraints. *Journal of Geometry and Physics*, 18(4):295–325, 1996.

[15] J. Cortes and M. de Leon. Reduction and reconstruction of the dynamics of nonholonomic systems. *Journal of Physics A-Mathematical and General*, 32(49):8615–8645, 1999.

[16] R.L. Devaney. Reversible diffeomorphisms and flows. *Trans. Am. Math. Soc.*, 218:89–113, 1976.

[17] A.T. Grigoryan and B.N. Fradlin. *A History of Mechanics of the Rigid Body.* Nauka, Moscow, 1982. Russian.

[18] J.K. Hale. *Ordinary Differential Equations, Pure and Applied Mathematics*, volume 21. Wiley Interscience, New York, 1969.

[19] A Iserles. Numerical analysis in lie groups. In *Foundations of computational mathematics*, pages 105–123. Oxford, 1999.

[20] W.S. Koon and J.E. Marsden. Poisson reduction for nonholonomic mechanical systems with symmetry. *Reports on mathematical physics*, 42(1-2):101–134, 1998.

[21] R. Kutteh. New approaches for molecular dynamics simulations with nonholonomic constraints. *Computer Physics Communications*, 119(2-3):159–168, 1999.

[22] J.L. Lagrange. *Méchanique Analytique.* Chez la Veuve Desaint, 1788.

[23] J.S.W. Lamb. Reversing symmetries in dynamical systems. *J. Phys. A*, 25:925–937, 1992.

[24] J.S.W. Lamb and J.A.G. Roberts. Time-reversal symmetry in dynamical systems. *Physica D. Nonlinear Phenomena*, pages 1–39, 1998.

[25] J. Loschmidt. *Wien. Ber.*, 75:67, 1877.

[26] J.E. Marsden and T.S. Ratiu. *Introduction to Mechanics and Symmetry.* Springer-Verlag New York, Inc., 1994.

[27] R.I. McLachlan and M. Perlmutter. Integrators for nonholonomic mechanical systems. Preprint, 2001.

[28] A. Nordmark and H. Essén. Systems with a preferred spin direction. *Proc. R. Soc. Lond. A*, 455:933–941, 1999.

[29] R. Penrose. *General Relativity*, chapter Singularities and time-asymmetry, pages 581–638. Cambridge Univ. Press, Cambridge, 1979.

[30] R. Penrose. *The Emperor's New Mind.* Oxford Univ. Press, Oxford, 1989.

[31] J.A.G. Roberts and G.R.W. Quispel. Chaos and time-reversal symmetry. order and chaos in reversible dynamical systems. *Physics Reports*, 216(2 & 3):66–110, July 1992.

[32] C. Robinson. *Dynamical Systems: stability, symbolic dynamics and chaos*, pages 146, 326. CRC Press, 1995.

[33] W. Sarlet, F. Cantrijn, and D.J. Saunders. A geometrical framework for the study of nonholonomic lagrangian systems. *Journal of Physics A-Mathematical and General*, 28(11):3253–3268, 1995.

[34] J.L. Synge. On geometry of dynamics. *Philos. Trans. R. Soc. Lond.*, 226(2):31–107, 1927.

[35] J.L. Synge. Geodesics in non-holonomic geometry. *Math. Ann.*, 99:738–751, 1928.

[36] A.M. Vershik and V.Ya. Gershkovich. Nonholonomic dynamical systems, geometry of distributions and variational problems. In *Dynamical systems. VII. Integrable systems, nonholonomic dynamical systems*, chapter 1-3, pages 4–81. Springer-Verlag, 1994.

[37] G. Vranceanu. Parallèlisme et courbure dans une variété non holonome. In *Atti del Congresso Internaz. dei Matematici*, volume 5, pages 63–75, Zanichelli: Bologna, 1931.

[38] E.P. Wigner. *Time Inversion, in: Group Theory*, chapter 26. Academic Press, New York, 1959.