THE UNIVERSITY OF CALGARY

A Family of Nonholonomic Systems with Symmetry

by

Creighton Garth MacDonnell

A DISSERTATION

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

May, 1997

0-612-24548-9

Table of Contents

List of Figures

Chapter 1

Introduction

In this introduction, we describe the process by which this dissertation was conceived, and discuss literature that played a role.

Reduction

We began the work leading to this dissertation by learning the theory of reduction in Hamiltonian systems, as described in the text by Arnold [2], and by Abraham and Marsden [1].

These texts (particularly [1]) employed the methods of differential geometry, with a level of abstraction that caused some uneasiness to this writer, who had learned mechanics from classical texts, such as Whittaker [34] and Goldstein [17]. The frustration arose because, although the mathematics was elegant, it was difficult to find correspondences between the mathematical viewpoint and "physical" viewpoint.

Nonholonomic Reduction, Quasi-Velocities

The work then turned to understanding recent attempts to develop a general theory of reduction for nonholonomic systems.

Loosely speaking, a nonholonomic mechanical system is one in which there are linear constraints among the "velocities", which cannot be removed by choosing coordinates for some subspace of the full configuration space. The question of what the equations of motion are, for such systems, is of some physical interest. Whittaker [34, article 87] attributes the extension of Lagrange's equations to nonholonomic systems to Ferrers [14, 1871], Neumann [26, 1888] and Vierkandt [32, 1892].

The paper by Bates and Śniatycki [5] proposed a reduction scheme for nonholonomic systems using the Hamiltonian formulation. Once again, we found it difficult to reconcile the mathematical abstraction with a physical viewpoint.

The paper by Koiller [25] used the Lagrangian formulation, to present a reduction scheme for *non-Abelian Chaplygin systems* (a subclass of nonholonomic systems).

Koiller's paper also used the concept of quasi-velocities. The notion of quasi-velocities is a fundamental tool in the theoretical discussion in this dissertation. Koiller indicates that quasi-velocities were advocated strongly by Hamel. Whittaker [34, article 30] gives credit for first finding "the Lagrangian equations for quasi-coordinates" to Boltzmann [8, 1902] and to Hamel [18, 1904]. Whittaker's description of *quasi-coordinates* and the associated equations of motion is not easily recognized as being the same concept as that used in this dissertation. Arnold, Kozlov and Nieshtadt [3, section 2.4 of chapter 1] give a more modern description of quasi-velocities. These authors give credit for finding the equations of motion for quasi-velocities to Poincaré, citing a date of 1901. In this dissertation, we call the equations of motion for quasi-velocities *Poincaré's equations*.

The paper by Bloch, Krishnaprasad, Marsden and Murray [7] also used a Lagrangian formulation, to discuss reduction and related ideas. The reduction procedure given there is essentially the same as the procedure appearing in Koiller, but the discussion uses a more geometric language to achieve greater conciseness. This paper also introduces the subspaces S and H, of the tangent space, which are used extensively in this dissertation (see chapter 6). Bloch et al. use this decomposition to obtain what they call the *momentum equation*. This momentum equation is closely related to the *adjoint* equation used in this dissertation.

Examples of Nonholonomic Systems

Another aspect our work was to look at specific examples of nonholonomic systems, with a view to seeing whether these reduction procedures could be applied to them. In particular, we were directed by our advisors towards the "Chaplygin sphere", the "tippe top" and the "rolling disk".

The Chaplygin Sphere

Arnold, Kozlov and Nieshtadt [3, section 1.2 of chapter 3 and section 4.1 of chapter 4] make some comments on the Chaplygin sphere. They do not however explain the solution of this problem in detail. A complete solution is given in Chaplygin [10], which unfortunately is written in Russian.

The Tippe Top

Many papers were written about the tippe top in the years from 1941 to 1994, including Fokker [15, 1941], Braams [9, 1952], Hugenholtz [20, 1952], Pliskin [30, 1954], Parkyn [28, 1958], Cohen [11, 1977], Kane and Levinson [23, 1978] and Or [27, 1994]. We could have listed several more. The tippe top appears to have spawned a small industry.

All of these authors, with the exception of Kane and Levinson, were apparently unaware of the analysis in Routh [31, article 243, 1905], which showed definitively that the rising behaviour could not be explained by a model with rolling without slipping. They were also apparently unaware of the paper by Gallop [16, 1903], which showed that the behaviour could be explained as a consequence of dissipation of energy by allowing slipping.

As we are concerned in this dissertation with nonholonomic systems (which would require rolling without slipping), we will not discuss the tippe top at any length. For a modern (and more exhaustive) treatment, we refer the reader to Ebenfeld and Scheck [13].

The Rolling Disk

It was shown by Kemppainen [24] that the reduction procedure of Bates and Śniatycki could be applied to the problem of the rolling disk. However, the procedure did not reduce the system to the fullest extent possible. That reduction procedure leads to a system on a 4-dimensional manifold.

In Pars [29, Section 8.12] and Routh [31, Article 244a], it was shown that this system could be reduced to an ODE involving θ and its time derivative $\dot{\theta}$, where θ is the angle between the vertical, and the line perpendicular to the disk, through its center (see figure 9.1 on page 72 and the examples following). This ODE involved the solution functions of another linear ODE system. This latter system gave rise to two constants of the motion, which were linear in the velocities.

Kemppainen (in effect) observed that the $(\theta, \dot{\theta})$-system, obtained for each value of the constants, was a 1-degree of freedom Hamiltonian system. From Kemppainen's point of view, the 4-dimensional "reduced space" was foliated by a "2-parameter family of 1-degree of freedom Hamiltonian systems". This point of view is also used in Cushman, Hermans and Kemppainen [12].

Symmetric Sphere on a Surface of Revolution

At about the same time, Hermans [19] was examining the geometry of a symmetric sphere rolling on a surface of revolution (with the central axis vertical).

This led us to the classical analysis of a symmetric sphere rolling on a surface of revolution (Routh [31, Article 230]). This analysis is strikingly similar to the analysis of the rolling disk. The angle θ is replaced in Routh's treatment by the angle between the vertical and the normal to the surface. As this angle may not uniquely identify a latitude of the surface, we prefer to use another, more robust quantity (see figure 10.1 on page 110), in our presentation.

Axially Symmetric Body On a Horizontal Plane

We were also led by Hermans to the classical analysis of a body of revolution rolling on a horizontal plane (Gallop [16, page 371] or Routh [31, Article 241a]).

The rolling disk is in fact just a special case of a body of revolution, as is the tippe top. The body of revolution may be solved (at least in the formal sense) by the same approach as the rolling disk. The line through the center of the disk is replaced by the axis of symmetry of the body.

The Direction Taken by This Dissertation

Several questions now arise:

- Is the reduced system, produced by these analyses, Hamiltonian in general, as it is for the rolling disk?
- What underlying geometric structure might be present in each of these two problems, and others, and led to this reduction?

- In conventional mechanical systems, constants of the motion can typically be associated, using Noether's theorem, with the momentum generated by some "action" on the configuration space. Is it possible to give a similar interpretation for the constants of the motion that arise in these analyses?

The consideration of these questions gave rise early on to the paper by Bates, Graumann and MacDonnell [4], and now, to this dissertation.

The Examples

In addition, we provide several specific examples of mechanical systems. In these examples, we use a somewhat unique approach to doing the calculations.

We believe our method makes it easier to avoid errors commonly made in the formulation of problems, by avoiding the need to consider fictional forces (easy to neglect) and to resolve forces into various components. Our approach also permits calculations to be advanced without the need to inject arguments involving "vector diagrams" (also prone to errors).

After an initial excursion into mathematical abstraction, we have attempted to bring a physical viewpoint to the examples.

Contributions to Knowledege by this Dissertation

As is common with doctoral dissertations, some final amendments and additions were made to this dissertation, at the request of the oral examination committee. This subsection is one such addition.

Much of this thesis is concerned with the analysis of an axially symmetric body rolling on a flat plane, and of a balanced sphere rolling on a surface of revolution about the vertical. Each of these problems has "classical" solutions.

In particular, the essential features of the reduced system for the axially symmetric body were known to Routh[31, article 243, 1905] and Gallop [16, 1903]. However, this knowledge was largely forgotten by 1950, as evidenced by the plethora of papers on the tippe top listed earlier.

To begin then, the author feels that the act of restoring this knowledge to currency has some value in itself.

Moreover, this knowledge has not simply been regurgitated from classical sources. This dissertation *invents* a modern method of reduction. This modern theory uses the concepts of Lie groups and differential geometry, which were not used classically. This theory is then applied to the axial symmetry problem and to the surface of revolution problem.

The recasting of the problems of the Lagrange top and the free body in a similiar framework, in recent decades, has certainly been regarded as having value. There is every reason to think that this should also be true for the problems dealt with here.

In addition, these treatments of the Lagrange top and free body used reduction in the Hamiltonian framework. This dissertation uses a method of reduction in the Lagrangian framework, invented for this purpose. The notions of the *flatness conditions* and the *adjoint equation* which are fundamental to this Lagrangian reduction method are new, to the best knowledge of the author.

At the time the author was beginning this work, several persons interested in nonholonomic mechanical systems had observed the striking similarity between the classical solutions for the rolling disk and for the sphere on a surface of revolution. Some felt it was only a coincidence. This dissertation, using its Lagrangian reduction procedure, has revealed clearly the common geometric structure, shared by these systems, that results in their reducibility, as well as that of the more general axially symmetric body problem.

The Lagrangian reduction procedure put forward in this dissertation was invented with the axially symmetric body and surface of revolution problems in mind. It was not anticipated that it would have application to the Chaplygin sphere problem. However it does, and allows one to conclude that this system may be reduced to a Hamiltonian system on the tangent bundle of the surface of a sphere (TS^2). This result is significant, because the existence of a constant of the motion for this reduced problem, independent of the energy, allows one to conclude that the Chaplygin sphere problem is integrable by quadratures. To the best knowledge of the author, this is a new result. Arnold, Kozlov and Nieshtadt [3, section 4.1 of chapter 4], discuss the integrability of this problem, with no mention of this fact.

Finally, the detailed analysis of the behaviour of a basketball also appears to be new. The merit of the approach used in this dissertation, is that a rigorous method had been used to obtain the reduced system, so that in the end, only a single first order system of ordinary differential equations needs to be solved numerically. The author is inclined to regard conclusions drawn from this approach as much more reliable than an approach involving the numerical solution of the full equations of motion. Certainly the rigorous reduction of the problem provides insight not obtainable from a purely numerical method.

Chapter 2

Physical Context and Background

In this section we provide a review of the physics motivating the mathematical structures discussed in this thesis.

2.1 Gravitational and Electromagnetic Forces

We will deal with mechanical systems that involve rigid bodies. However, to simplify matters in the beginning, we will first consider a system consisting of point particles, each with mass and possibly electric charge.

We suppose there to be N particles. The i-th particle has mass m_i, and electric charge q_i. We take a set of right-handed cartesian axes, and suppose the coordinates of the i-th particle with respect to these axes to be $x_i = (x_{i1}, x_{i2}, x_{i3})$.

We suppose there to be *static* gravitational, electric, and magnetic fields present. The assumption that these fields are static is of course an idealization. But this assumption is commonly made in posing problems in mechanics.

Under these assumptions, we may suppose that there exists a scalar gravitational potential ϕ, a scalar electric potential ψ and a vector magnetic field $B = (B_1, B_2, B_3)$. If we define

$$U(x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23}, \dots) = \sum_i \{m_i\phi(x_{i1}, x_{i2}, x_{i3}) + q_i\psi(x_{i1}, x_{i2}, x_{i3})\}$$

and introduce the 3-index permutation symbol, ε_{jkm} then the j-th component P_{ij} of the force P_i on the i-th particle due to gravity and the electro-magnetic field is given by

$$P_{ij} = -\frac{\partial U}{\partial x_{ij}} + \sum_{k,m} \varepsilon_{jkm} \dot{x}_{ik} B_m(x_{i1}, x_{i2}, x_{i3}).$$

2.2 The Lagrangian and Generalized Forces

If we use F_i to represent any other force on the i-th particle, Newton's second law becomes

$$(2.1) \qquad m_i \ddot{x}_{ij} = -\frac{\partial U}{\partial x_{ij}} + \sum_{k,m} \varepsilon_{jkm} \dot{x}_{ik} B_m(x_{i1}, x_{i2}, x_{i3}) + F_{ij},$$

where '$\cdot$' indicates differentiation with respect to time. Define the *kinetic energy* K by

$$K(x, \dot{x}) = \frac{1}{2} \sum_i m_i \dot{x}_i^T \dot{x}_i = \frac{1}{2} \sum_{i,j} m_i \dot{x}_{ij}^2,$$

and then the *Lagrangian* by

$$L(x, \dot{x}) = K(x, \dot{x}) - U(x).$$

Then equation 2.1 may be rewritten

$$\left(\frac{\partial L}{\partial \dot{x}_{ij}}\right)' - \frac{\partial L}{\partial x_{ij}} - \sum_{k,m} \varepsilon_{jkm} \dot{x}_{ik} B_m(x_{i1}, x_{i2}, x_{i3}) = F_{ij},$$

where '$'$' denotes differentiation with respect to time.

Now suppose that we wish to use an alternative set of *generalized coordinates* $\{q_1, \ldots, q_{3N}\}$ to describe the configuration of the system. Then a straight forward calculation will yield

$$(2.2) \qquad \left(\frac{\partial L}{\partial \dot{q}_i}\right)' - \frac{\partial L}{\partial q_i} - \sum_j M_{ij} \dot{q}_j = Q_i,$$

where

$$Q_i = \sum_{j,k} F_{jk} \frac{\partial x_{jk}}{\partial q_i}$$

is called the *generalized force*, and

$$M_{ij} = \sum_{k,m,r,s} \varepsilon_{mrs} B_s(x_{k1}, x_{k2}, x_{k3}) \frac{\partial x_{km}}{\partial q_i} \frac{\partial x_{kr}}{\partial q_j}$$

is called the *generalized magnetic field*. If we make another change of coordinates to the set $\{\overline{q}_1, \ldots, \overline{q}_{3N}\}$, the equation of motion will become

$$\left(\frac{\partial L}{\partial \dot{\overline{q}}_i}\right)' - \frac{\partial L}{\partial \overline{q}_i} - \sum_j \overline{M}_{ij} \dot{\overline{q}}_j = \overline{Q}_i,$$

with

$$\overline{Q}_i = \sum_j Q_j \frac{\partial q_j}{\partial \overline{q}_i}$$

and

$$\overline{M}_{ij} = \sum_{k,m} M_{km} \frac{\partial q_k}{\partial \overline{q}_i} \frac{\partial q_m}{\partial \overline{q}_j}.$$

Equations 2.2 are called *Lagrange's equations*.

Whatever coordinates we use for the system, the Lagrangian will have the form $L = K - U$, while the kinetic energy will take the general form

$$K(q, \dot{q}) = \frac{1}{2} \sum_{i,j} K_{ij}(q) \dot{q}_i \dot{q}_j.$$

2.3 Holonomic Constraints and Associated Forces

Now suppose that there are constraints of the form

$$G_i(t, q) = 0 \tag{2.3}$$

for the system, where t represents time. We will call such constraints *holonomic*. The constraints could for example require that the distance between any two particles remains fixed, or that a particle is constrained to remain on a surface.

In the presence of such a constraint, there must be forces included in the generalized forces appearing in equation 2.2, which serve solely to enforce these constraints. We will call these *forces of constraint*. In order to distinguish these forces, we will rewrite equation 2.2 as

$$\left(\frac{\partial L}{\partial \dot{q}_i}\right)' - \frac{\partial L}{\partial q_i} - \sum_j M_{ij}\dot{q}_j = R_i + Q_i.$$

in which R_i represents the forces of constraint, and Q_i represents other *external* forces, not derived from the potentials already embodied in L.

Suppose now that the solution manifold to equation 2.3 can be parameterized locally as

$$q = q(t, \overline{q}).$$

Then by a calculation identical to the one used above for the change of coordinates, we may obtain

$$\left(\frac{\partial L}{\partial \dot{\overline{q}}_i}\right)' - \frac{\partial L}{\partial \overline{q}_i} - \sum_j \overline{M}_{ij}\dot{\overline{q}}_j = \sum_j R_j \frac{\partial q_j}{\partial \overline{q}_i} + \overline{Q}_i.$$

In typical problems to which Lagrange's Equations are applied, the external forces Q_i (and therefore $\overline{Q}_i$) are clearly zero, and there is no magnetic field. However, the term involving the constraint force is also normally assumed to be zero. This requires some justification.

The assumption that

$$\sum_j R_j \frac{\partial q_j}{\partial \overline{q}_i} = 0 \tag{2.4}$$

holds is called *D'Alembert's principle*. This principle is often expressed by a statement such as "the virtual work done by the forces of constraint is zero".

The physical rationale for D'Alembert's principle is illustrated by the following examples:

- point mass on a surface

  D'Alembert's principle in this context asserts that the forces that keep the particle on the surface do not push the particle around on the surface.

- rigid body on an axle

  The forces of constraint here hold the body together, and prevent the body from moving, except for rotating on the axle. D'Alembert's principle asserts that these forces will not make the body spin faster or slower about this axis.

Whittaker ([34, Article 25]) describes a *holonomic dynamical system* as one "for which a displacement represented by arbitrary infinitesimal changes in the coordinates is in general a possible displacement". When this condition is not satisfied, he refers to the system as *nonholonomic*.

This description is in fact ambiguous. In terms of the original q_i coordinates above, our hypothetical system does not satisfy Whittaker's condition. In terms of the $\overline{q}_i$ coordinates, it does.

Thus our hypothetical system may be regarded as a holonomic system only if we think of the configuration manifold as being the solution manifold of equation 2.3. Otherwise it is a nonholonomic system.

We discuss nonholonomic systems at greater length in the next subsection.

2.4 Nonholonomic Constraints and Associated Forces

In the previous section, we introduced constraints (equation 2.3) of the form

$$G_i(t,q) = 0,$$

and called these holonomic. In physical problems, it is possible to have constraints of the form

$$\sum_j B_{ij}(q)\dot{q}_j + b_i(q) = 0. \tag{2.5}$$

We will call constraints of this form *nonholonomic*. As for holonomic constraints, there must be forces of constraint present in Lagrange's equations to enforce the constraints. For holonomic constraints, we were able to find (in principle at least) an alternate set of coordinates, such that we could just ignore the forces of constraint. For nonholonomic constraints, there may not be such a set of coordinates.

It is helpful in this context to consider how we could handle the forces of constraint for holonomic constraints, without using local coordinates for the solution manifold.

If we differentiate our holonomic constraints with respect to time we obtain

$$\sum_j \frac{\partial G_i}{\partial q_j}\dot{q}_j + \frac{\partial G_i}{\partial t} = 0. \tag{2.6}$$

This has the same form as equation 2.5. Our approach will be to include these equations with Lagrange's equations, and to seek a solution of this larger system. The difficulty that remains, is that the (as yet unknown) forces of constraint appear in the equations. Our system of equations must permit the determination of these forces simultaneously with the determination of the trajectory.

Towards this end, observe that differentiating our original holonomic constraints with respect to local coordinates for the solution manifold gives

$$\sum_j \frac{\partial G_i}{\partial q_j}\frac{\partial q_j}{\partial \overline{q}_k} = 0.$$

At this point we will use some basic concepts from linear algebra. The equation above says that the row vectors

$$\begin{bmatrix} \frac{\partial G_i}{\partial q_1} & \frac{\partial G_i}{\partial q_2} & \cdots \end{bmatrix},$$

(one for each value of i), lie in the left null space of the matrix

$$\begin{bmatrix} \frac{\partial q_1}{\partial \overline{q}_1} & \frac{\partial q_1}{\partial \overline{q}_2} & \cdots \\ \frac{\partial q_2}{\partial \overline{q}_1} & \frac{\partial q_2}{\partial \overline{q}_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$

But these row vectors are linearly independent, and therefore span the left null space. D'Alembert's principle,

$$\sum_j R_j \frac{\partial q_j}{\partial \overline{q}_k} = 0,$$

says that the row vector

$$\begin{bmatrix} R_1 & R_2 & \cdots \end{bmatrix}$$

also lies in the left null space. It follows that there is a set of multipliers $\{\lambda_1, \lambda_2, \dots\}$ such that

$$R_i = \sum_j \lambda_j \frac{\partial G_j}{\partial q_i}$$

holds. With this result, our system of equations becomes

$$\begin{aligned}
\left(\frac{\partial L}{\partial \dot{q}_i}\right)' - \frac{\partial L}{\partial q_i} - \sum_j M_{ij}\dot{q}_j &= \sum_j \lambda_j \frac{\partial G_j}{\partial q_i} \\
\sum_k \frac{\partial G_j}{\partial q_k}\dot{q}_k + \frac{\partial G_j}{\partial t} &= 0.
\end{aligned}$$

Solving this system involves finding both q and λ as functions of time.

We now return to the question of the forces of constraint needed to enforce the nonholonomic constraints in equation 2.5. A comparison of equation 2.5 with equation 2.6 leads one to suspect that these forces should satisfy

$$R_i = \sum_j \lambda_j B_{ji}.$$

There is also a physical rationale for making this assumption. The constraint provided by equation 2.5, or equation 2.6, is that at each point q in the configuration manifold, there is a plane in the tangent space to which $\dot{q}$ is restricted. It appears natural to suppose that the forces of constraint should depend only on q, $\dot{q}$ and the plane at q. They should not depend on the plane at some other point. This is consistent with the local quality of physical laws, which appears to be quite general in physics.

The system of equations we are ultimately led to is

$$\begin{aligned}
\left(\frac{\partial L}{\partial \dot{q}_i}\right)' - \frac{\partial L}{\partial q_i} - \sum_j M_{ij}\dot{q}_j &= \sum_j \lambda_j B_{ji} \\
\sum_k B_{jk}\dot{q}_k + b_j &= 0.
\end{aligned} \tag{2.7}$$

In practise, these equations seem to lead to physically reasonable equations of motion. Some other schemes do not (see for example the discussion of *vakonomic mechanics* in [3]).

Whittaker ([34, Article 87]) calls these equations *Lagrange's equations with undetermined multipliers*, and attributes the extension of Lagrange's equations to nonholonomic systems to Ferrers, citing a paper from 1871 ([14]), Neumann, citing a paper from 1888 ([26]), and Vierkandt, citing a paper from 1892 ([32]).

2.5 The Kinematics of a Rigid Body

In this section we review the kinematics of rigid bodies, and seek to familiarize the reader with the notation which will be used in our examples.

We will use the notation of linear algebra. A vector will normally be represented as a *column vector*. We will regard $\Re^3$ as containing column vectors. We represent the identity matrix by I. For the standard unit column vectors we use

$$\epsilon_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \epsilon_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \epsilon_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

We will frequently use the map $A : \Re^3 \to \mathfrak{so}(3)$ defined by

$$A\left(\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}\right) = \begin{bmatrix} 0 & -z_3 & z_2 \\ z_3 & 0 & -z_1 \\ -z_2 & z_1 & 0 \end{bmatrix}.$$

The expression $A(y)z$ will thus be equivalent to taking the cross product of y and z. The properties

$$\begin{aligned} A(y)z &= -A(z)y, \\ A(Wz) &= WA(z)W^T, \ \forall W \in SO(3), \\ A(y)A(z) &= zy^T - (y^Tz)I \qquad \text{and} \\ A(A(y)z) &= A(y)A(z) - A(z)A(y) \end{aligned}$$

hold. We will also use the map $R : \Re^3 \to SO(3)$ defined by

$$R(z) = \exp A(z).$$

For $z \in \Re^3$ and $\phi \in \Re$, observe that

$$\frac{\partial R(\phi z)}{\partial \phi} = A(z)R(\phi z) = R(\phi z)A(z)$$

holds.

For $z \in \Re^3$ a *unit* vector and $\phi \in \Re$, $R(\phi z)$ corresponds to a rotation through an angle of magnitude ϕ, about the direction given by z (using the "right-handed rule"). We will also use the short hand notation

$$R_n(\phi) = R(\phi e_n).$$

We assume there to be a set of cartesian axes fixed in space. Associated with these fixed axes we have a set of orthonormal vectors, given by the standard unit column vectors with respect to the fixed axes.

We now consider a rigid body. Let the position of the center of mass be given with respect to the fixed axes by

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 e_1 + x_2 e_2 + x_3 e_3.$$

We assume there to be another set of cartesian axes fixed in the body, with the origin lying at the center of mass. Associated with these moving axes we have a corresponding moving set of orthonormal vectors, given by w_1, w_2 and w_3 with respect to the fixed axes.

The orientation of the body is given by the special orthogonal matrix $W \in SO(3)$ with w_1, w_2 and w_3 as columns,

$$W = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}.$$

Denote the *angular velocity* vector of the body with respect to the fixed axes by ν, and with respect to the body (moving) axes by ω. These vectors are determined by

$$\dot{W} = WA(\omega) = A(\nu)W,$$

and will always be related by

$$\nu = W\omega.$$

Suppose points in the body have positions given by a vector y with respect to the moving axes. Then the position of each point with respect to the fixed axes is given by the vector z determined by

$$z = x + Wy.$$

Differentiating with respect to time we get

$$\dot{z} = \dot{x} + W A(\omega) y.$$

or alternatively

$$\begin{aligned} \dot{z} &= \dot{x} + A(\nu) W y \\ &= \dot{x} + A(\nu)\{z - x\}. \end{aligned} \tag{2.8}$$

Let the total mass of the body be m. Let $\rho = \rho(y)$ give the density of the body. Then

$$\begin{aligned} \int dy^3 \rho(y) &= m \qquad \text{and} \\ \int dy^3 y \rho(y) &= 0 \end{aligned}$$

hold.

We may now calculate, for the kinetic energy of the body,

$$\begin{aligned} K &= \frac{1}{2} \int dy^3 \rho(y) \dot{z}(y)^T \dot{z}(y) \\ &= \frac{1}{2} \int dy^3 \rho(y) [\dot{x} + W A(\omega) y]^T [\dot{x} + W A(\omega) y] \\ &= \frac{1}{2} m \dot{x}^T \dot{x} + \frac{1}{2} \omega^T J \omega \end{aligned} \tag{2.9}$$

with

$$\begin{aligned} J &= -\int dy^3 \rho(y) A(y) A(y) \\ &= \int dy^3 \rho(y) [(y^T y) I - y y^T]. \end{aligned}$$

J is the *inertia matrix* of the body with respect to the body axes. J is constant in time, and symmetric. By choosing the body axes to be the principal axes of J, we may take J to be diagonal.

$$J = \begin{bmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_3 \end{bmatrix}.$$

where J_1, J_2 and J_3 are the *principal moments of inertia* of the body.

Another quantity often used in mechanics, is the *angular momentum* with respect to a point in space. For simplicity, suppose the point to be the origin. The angular momentum is

$$
\begin{aligned}
P &= \int dy^3 \rho(y) A(z(y)) \dot{z}(y) \\
&= \int dy^3 \rho(y) A(x + Wy)[\dot{x} + WA(\omega)y] \\
&= mA(x)\dot{x} + \int dy^3 \rho(y) WA(y)A(\omega)y] \\
&= mA(x)\dot{x} + WJ\omega. \qquad (2.10)
\end{aligned}
$$

The second term in this expression, $WJ\omega$, which does not depend on the choice of reference point, is regarded as the angular momentum of the body, *in spatial coordinates*. The expression $J\omega$ is regarded as the angular momentum of the body *in body coordinates*. The first term we regard as the angular momentum of the center of mass.

Chapter 3

Basic Mathematical Definitions

In this section, we use the ideas of the previous section to motivate a more precise mathematical framework. All manifolds will be assumed to be of finite dimension.

3.1 Preliminary Definitions

Definition 3.1.1. *Let Q be a manifold, and $L : TQ \to \Re$ a function. The* Legendre transformation *of L is the function $FL : TQ \to T^*Q$ given by*

$$\langle FL(\tau), \rho\rangle = \left.\frac{d}{ds}\right|_{s=0} L(\tau + s\rho).$$

Observe that $FL : TQ \to T^*Q$ is *fibre preserving*. That is to say, if $\pi : TQ \to Q$ and $\pi^* : T^*Q \to Q$ are the respective bundle projections, then $\pi^* \circ FL = \pi$.

Definition 3.1.2. *Let Q be a manifold, and $L : TQ \to \Re$ a function. The* Hessian transformation *of L is the function $F^2L : TQ \to T_2^0Q$ given by*

$$\begin{aligned}
F^2L(\tau)(\rho, \sigma) &= \left.\frac{d}{du}\right|_{u=0} \langle FL(\tau + u\sigma), \rho\rangle \\
&= \left.\frac{\partial}{\partial u}\right|_{u=0} \left\{ \left.\frac{\partial}{\partial s}\right|_{s=0} L(\tau + u\sigma + s\rho)\right\} \\
&= \left.\frac{\partial^2}{\partial u \partial s}\right|_{u=0, s=0} L(\tau + u\sigma + s\rho).
\end{aligned}$$

Observe that $F^2L(\tau)$ is symmetric for each τ, and that $F^2L : TQ \to T^0_2Q$ is fibre preserving.

Definition 3.1.3. *Let Q be a manifold, and $L : TQ \to \mathbb{R}$ a function. Then L is a* Lagrangian *on Q if and only $F^2L(\tau)$ is positive definite for each τ.*

Many authors call any function $L : TQ \to \mathbb{R}$ a Lagrangian. What we have called a Lagrangian they refer to as a *regular Lagrangian.*

Definition 3.1.4. *A* nonholonomic system *is a 5-tuple (Q, L, M, D, β), where*

- *Q is a manifold.*
- *$L : TQ \to \mathbb{R}$ is a Lagrangian on Q.*
- *M is a 2-form on Q.*
- *D is a distribution on Q and*
- *β is a vector field on Q.*

We call M the *magnetic 2-form.*

We call D the *constraint distribution.* Associated with D we define $D_\beta = \{\tau \in TQ | \tau - \beta \circ \pi(\tau) \in D\}$.

Associated with any nonholonomic system there is a collection of systems of equations for a trajectory, $q = q(t)$ in Q, which we describe below. These equations will have the effect of restricting the trajectory to lie in D_β.

3.2 Lagrange's Equations

To describe a typical system of equations in the collection referred to in definition 3.1.4, let $q^i : V \to \mathbb{R}$ be local coordinates on Q. We will use the notation $\dot{q}^i$ for dq^i, as normally done in mechanics.

Also, let $\{B^1, \ldots, B^r\}$ be a set of independent 1-forms on V that locally determine D. That is to say, we have $D \cap TV = \{\tau \in TV | \langle B^i, \tau \rangle = 0, \forall i\}$. We may of course need to reduce the size of V in order to make such a choice.

Then define $M_{ij} = M(\partial/\partial q^i, \partial/\partial q^j)$, $B^i_j = \langle B^i, \partial/\partial q^j \rangle$ and $b^i = \langle B^i, \beta \rangle$.

The system of equations in question is (using the summation convention)

$$\left(\frac{\partial L}{\partial \dot{q}^i}\right)' - \frac{\partial L}{\partial q^i} - M_{ij}\dot{q}^j = \lambda_j B^j_i \tag{3.1}$$

$$B^j_i \dot{q}^i = b^j. \tag{3.2}$$

These are called *Lagrange's equations.*

In these equations. $\lambda_1, \ldots, \lambda_r$ are called *undetermined multipliers*. and are to be determined as functions of t along with the q^i.

Proposition 3.2.1. *For each $q_0 \in Q$. the collection of Lagrange's Equations for a nonholonomic system uniquely determines a trajectory $q = q(t)$ with $q(0) = q_0$.*

Note that we do **not** claim that $\lambda = \lambda(t)$ is uniquely determined.

Proof. We proceed in three steps.

1. First we show that in a local coordinate system. the λ variables may be eliminated. thereby leaving an ODE (Ordinary Differential Equation).

   Observe that in coordinates.

   $$F^2L(\tau) = \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}(\tau) dq^i \otimes dq^j$$

   holds. and so the matrix A. with coefficients given by

   $$A_{ij} = \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}.$$

   must be positive definite on TV. Then. from the calculation

   $$\left(\frac{\partial L}{\partial \dot{q}^i}\right)' = \frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i}\ddot{q}^j + \frac{\partial^2 L}{\partial q^j \partial \dot{q}^i}\dot{q}^j.$$

   we see that equation 3.1 has the general form

   $$A_{ij}\ddot{q}^j = \lambda_j B_i^j + f_i(q, \dot{q}).$$

   Differentiating equation 3.2 we obtain an equation with the general form

   $$B_i^j \ddot{q}^i = g^j(q, \dot{q}).$$

   Using matrix notation in an obvious manner. these two forms give us

   $$BA^{-1}B^T\lambda = -BA^{-1}f + g.$$

But the matrix $BA^{-1}B^T$ is non-singular, by the sequence of inferences

$$\begin{aligned} BA^{-1}B^T\lambda = 0 \quad &\Rightarrow \quad (A^{-1}B^T\lambda)^T A(A^{-1}B^T\lambda) = 0 \\ &\Rightarrow \quad A^{-1}B^T\lambda = 0 \\ &\Rightarrow \quad \lambda^T B = 0 \\ &\Rightarrow \quad \lambda = 0 \end{aligned}$$

Hence the λ variables may be expressed in terms of q and $\dot{q}$, so that we have a second order ODE for q.

2. Next we show that the trajectory found is independent of the choice of coordinates. To this end let $\overline{q}^i = \overline{q}^i(q)$ be a coordinate transformation. This induces a coordinate transformation

$$\dot{\overline{q}}^i(q,\dot{q}) = \frac{\partial \overline{q}^i}{\partial q^k}(q)\dot{q}^k$$

on TQ. Let $\overline{L}(\overline{q},\dot{\overline{q}})$ be the expression for the Lagrangian and $\overline{M}_{ij}(\overline{q})$ for the magnetic 2-form coefficients, in the new coordinates, so that we have

$$\begin{aligned} L(q,\dot{q}) &= \overline{L}\left(\overline{q}(q), \frac{\partial \overline{q}}{\partial q^k}(q)\dot{q}^k\right) \qquad \text{and} \\ M_{ij}(q) &= \overline{M}_{kr}(\overline{q}(q))\frac{\partial \overline{q}^k}{\partial q^i}(q)\frac{\partial \overline{q}^r}{\partial q^j}(q). \end{aligned}$$

A simple calculation then yields

$$\left(\frac{\partial L}{\partial \dot{q}^i}\right)' - \frac{\partial L}{\partial q^i} - M_{ij}\dot{q}^j = \left\{\left(\frac{\partial \overline{L}}{\partial \dot{\overline{q}}^k}\right)' - \frac{\partial \overline{L}}{\partial \overline{q}^k} - \overline{M}_{kj}\dot{\overline{q}}^j\right\}\frac{\partial \overline{q}^k}{\partial q^i}.$$

Since, in addition, we have

$$\begin{aligned} B_i^j &= \left\langle B^j, \frac{\partial}{\partial \overline{q}^k}\right\rangle \frac{\partial \overline{q}^k}{\partial q^i} \\ &= \overline{B}_k^j \frac{\partial \overline{q}^k}{\partial q^i}. \end{aligned}$$

we see that equations 3.1 and 3.2 hold in the new coordinates.

3. Finally we show that the trajectory found is independent of the choice of 1-forms $B^1, \ldots, B^r$. To this end let $\overline{B}^1, \ldots, \overline{B}^r$ be another choice.

   By reducing, if necessary, the size of the neighborhood V on which the coordinates and 1-forms are defined we may assume that there are functions $\alpha^j_k : V \to \mathbb{R}$, such that $B^j = \alpha^j_k \overline{B}^k$ on V holds, and such that the matrix P with coefficients given by $P_{jk} = \alpha^j_k$ is non-singular.

   We then have

$$B^j_i = \left\langle \alpha^j_k \overline{B}^k, \partial/\partial q^i \right\rangle = \alpha^j_k \overline{B}^k_i$$

   and

$$b^j = \left\langle \alpha^j_k \overline{B}^k, \beta \right\rangle = \alpha^j_k \overline{b}^k.$$

   We see now that equation 3.2 holds for the new 1-forms, using the same trajectory $q = q(t)$.

   Also, equation 3.1 holds with $q = q(t)$, if we replace λ_j by $\lambda_k \alpha^k_j$.

This completes the proof. □

Definition 3.2.1. *The* energy function *for L on TQ, which we denote $E : TQ \to \mathbb{R}$, is defined by*

$$E(\tau) = \langle FL(\tau), \tau \rangle - L(\tau).$$

For many important nonholonomic systems, $\beta \equiv 0$ holds. For such systems we have:

Proposition 3.2.2. *For the trajectory of a nonholonomic system with $\beta \equiv 0$, the energy remains constant. That is to say, $E \circ \dot{q}(t)$ is a constant function of t.*

Proof. Observe that in coordinates,

$$E(\tau) = \frac{\partial L}{\partial \dot{q}^i}(\tau)\dot{q}^i(\tau) - L(\tau)$$

holds, and so we may calculate

$$
\begin{aligned}
E' &= \left\{ \left(\frac{\partial L}{\partial \dot q^i} \right)' \dot q^i + \frac{\partial L}{\partial \dot q^i} \ddot q^i \right\} \\
&\quad - \left\{ \frac{\partial L}{\partial q^i} \dot q^i + \frac{\partial L}{\partial \dot q^i} \ddot q^i \right\} \\
&= \left\{ \left(\frac{\partial L}{\partial \dot q^i} \right)' - \frac{\partial L}{\partial q^i} \right\} \dot q^i \\
&= M_{ij} \dot q^i \dot q^j + \lambda_j B^j_i \dot q^i \\
&= \lambda_j \left\langle B^j, \beta \right\rangle = 0.
\end{aligned}
$$

□

3.3 Intrinsic Form of Lagrange's Equations

It is possible to formulate the equations for the trajectory of a nonholonomic system entirely in terms of globally defined geometric objects. We include the basic ideas used in this approach here for the sake of completeness.

Let $\pi^T : T(TQ) \to TQ$ be the tangent bundle projection map. Also, recall that from $\pi : TQ \to Q$ we may form the tangent map $T\pi : T(TQ) \to TQ$.

Recall too that we may use π to pull the magnetic 2-form M on Q back to a 2-form $\pi^* M$ on $T^*(TQ)$ (that is, $(\pi^* M)(\xi, \zeta) = M(T\pi(\xi), T\pi(\zeta))$).

Definition 3.3.1. *The* canonical 1-form *for L on TQ, which we denote θ_L, is defined by*

$$\langle \theta_L, \xi \rangle = \left\langle FL \circ \pi^T(\xi), T\pi(\xi) \right\rangle .$$

Definition 3.3.2. *The* constraint co-distribution *on TQ, which we denote D^0, is defined by*

$$D^0 = \{ \alpha \in T^*(TQ) | \langle \alpha, \xi \rangle = 0, \forall \xi \in (T\pi)^{-1}(D) \}.$$

Definition 3.3.3. *The* intrinsic form of Lagrange's equations *consists of the two requirements*

$$\ddot q(t) \lrcorner \left(d\theta_L + \frac{1}{2} \pi^* M \right) + dE \in D^0$$

and

$$\dot{q}(t) \in D_3.$$

In these expressions. $E : TQ \to \Re$ is the *energy function* introduced earlier. Again. this function is given by

$$E(\tau) = \langle FL(\tau).\tau \rangle - L(\tau).$$

Proposition 3.3.1. *Lagrange's Equations are a local consequence of the intrinsic form of Lagrange's Equations.*

Proof. It is obvious that the second intrinsic condition leads to equation 3.2. It remains to obtain equation 3.1.

For the purposes of this proof. we will denote the local coordinates on Q by z^i. We then introduce coordinates on TQ by $x^i = z^i \circ \pi$ and $v^i = dz^i$.

We now introduce the functions p_i defined by

$$p_i(\tau) = \frac{\partial L}{\partial v^i}(\tau).$$

from which we have

$$FL(\tau) = p_i(\tau)\, dz^i.$$

This leads immediately to

$$E(\tau) = p_i(\tau)\, v^i(\tau) - L(\tau).$$

giving

$$dE = v^i\, dp_i - \frac{\partial L}{\partial x^i}\, dx^i.$$

Next note that we have

$$\begin{aligned} \left\langle \theta_L . \frac{\partial}{\partial v^i} \right\rangle &= \left\langle FL \circ \pi^T \left(\frac{\partial}{\partial v^i} \right) . T\pi \left(\frac{\partial}{\partial v^i} \right) \right\rangle \\ &= 0. \end{aligned}$$

since

$$T\pi \left(\frac{\partial}{\partial v^i} \right) = 0$$

holds. We also have

$$\begin{aligned}\left\langle \theta_L, \frac{\partial}{\partial x^i} \right\rangle &= \left\langle FL \circ \pi^T \left(\frac{\partial}{\partial x^i}\right), T\pi \left(\frac{\partial}{\partial x^i}\right) \right\rangle \\ &= \left\langle p_j \, dz^j, \frac{\partial}{\partial z^i} \right\rangle \\ &= p_i.\end{aligned}$$

Thus we obtain

$$\theta_L = p_i \, dx^i.$$

From this we calculate

$$d\theta_L = dp_i \wedge dx^i.$$

In a similiar fashion, if we introduce the functions M_{ij} by

$$M_{ij} = M\left(\frac{\partial}{\partial z_i}, \frac{\partial}{\partial z_j}\right).$$

we have

$$\pi^* M = (M_{ij} \circ \pi) dx^i \wedge dx^j.$$

We may now calculate

$$\begin{aligned}&\ddot{q}(t) \lrcorner \left(d\theta_L + \frac{1}{2}\pi^* M\right) + dE \\ &\quad = \{\ddot{q}(t) \lrcorner dp_i\} \, dx^i - \{\ddot{q}(t) \lrcorner dx^i\} \, dp_i \\ &\qquad + \frac{1}{2} M_{ij}(q(t)) \{[\ddot{q}(t) \lrcorner dx^i] dx^j - [\ddot{q}(t) \lrcorner dx^j] dx^i\} \\ &\qquad\quad + v^i(t) \, dp_i - \frac{\partial L}{\partial x^i}(\dot{q}(t)) \, dx^i \\ &\quad = \left\{ \dot{p}_i(t) - \frac{\partial L}{\partial x^i}(\dot{q}(t)) - M_{ij}(q(t)) \dot{x}^j \right\} dx^i \\ &\qquad\quad + \{v^i(t) - \dot{x}^i(t)\} \, dp_i \\ &\quad = \left\{ \left(\frac{\partial L}{\partial v^i}\right)' - \frac{\partial L}{\partial x^i} - M_{ij} \dot{x}^j \right\} dx^i + \{v^i - \dot{x}^i\} \, dp_i.\end{aligned}$$

This last expression must be in D^0. But $\langle o, \partial/\partial v^j \rangle, \forall o \in D^0$ holds, so we must have

$$
\begin{aligned}
0 &= \{v^i - \dot{x}^i\} \frac{\partial p_i}{\partial v^j} \\
&= \{v^i - \dot{x}^i\} \frac{\partial^2 L}{\partial v^j \partial v^i}.
\end{aligned}
$$

Thus $v^i = \dot{x}^i$ must hold.

Now let $\{B^1, \ldots, B^r\}$ be a set of independent 1-forms locally determining D. Then $\{\pi^* B^1, \ldots, \pi^* B^r\}$ is a local basis for D^0. So for some undetermined multipliers $\lambda_1, \ldots, \lambda_r$, we must have

$$
\begin{aligned}
\left\{ \left(\frac{\partial L}{\partial v^i} \right)' - \frac{\partial L}{\partial x^i} - M_{ij} x^j \right\} dx^i &= \lambda_j \, \pi^* B^j \\
&= \lambda_j \, \pi^* (B^j_i \, dz^i) \\
&= \lambda_j \, B^j_i \, dx^i.
\end{aligned}
$$

The desired result now follows. □

3.4 Hamilton's Principle

In this subsection we show that the problem of finding a trajectory for a nonholonomic system is equivalent to solving a variational problem. What we describe here is not quite what is usually referred to as Hamilton's Principle, but is closely related.

Suppose we have a trajectory $q = q(t)$. Choose times t_1 and t_2. Consider the integral $\int_{t_1}^{t_2} L \, dt$. Supposing for the moment that the trajectory remains in one coordinate neighborhood between t_1 and t_2, we may calculate the effect of an infinitesimal variation of this trajectory on the integral:

$$
\begin{aligned}
\delta \int_{t_1}^{t_2} L \, dt &= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q^i} \delta q^i + \frac{\partial L}{\partial \dot{q}^i} \delta \left(\frac{dq^i}{dt} \right) \right) dt \\
&= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q^i} \delta q^i + \frac{\partial L}{\partial \dot{q}^i} \frac{d(\delta q^i)}{dt} \right) dt \\
&= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q^i} - \left(\frac{\partial L}{\partial \dot{q}^i} \right)' \right) \delta q^i \, dt + \left[\frac{\partial L}{\partial \dot{q}^i} \delta q^i \right]_{t_1}^{t_2} \\
&= \left[\frac{\partial L}{\partial \dot{q}^i} \delta q^i \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} M_{ij} \dot{q}^i \delta q^j \, dt - \int_{t_1}^{t_2} \lambda_j B^j_i \delta q^i \, dt
\end{aligned}
$$

Observing that in coordinates,

$$\begin{aligned} FL &= \frac{\partial L}{\partial \dot{q}^i} dq^i \\ &= \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i \end{aligned}$$

holds, we may write our result as

$$\delta \int_{t_1}^{t_2} L\,dt + \int_{t_1}^{t_2} M(\dot{q}, \delta q)dt = \langle FL, \delta q\rangle |_{t_1}^{t_2} - \int_{t_1}^{t_2} \lambda_j B_i^j \delta q^i\,dt.$$

If we restrict the variation to $\delta q \in D$, we have

$$\delta \int_{t_1}^{t_2} L\,dt + \int_{t_1}^{t_2} M(\dot{q}, \delta q)dt = \langle FL, \delta q\rangle |_{t_1}^{t_2} \tag{3.3}$$

This result is now coordinate independent, and may be extended beyond a single coordinate neighborhood.

The result following from this calculation is:

Proposition 3.4.1. *For a trajectory $q = q(t)$, the requirement that the trajectory satisfy Lagrange's Equations is equivalent to requiring that:*

- *The trajectory satisfies equation 3.3 for all values of t_1 and t_2 in its domain, when the variation of the trajectory satisfies $\delta q \in D$.*
- *The condition $dq/dt \in D_3$ holds.*

3.5 Natural Lagrangians

Definition 3.5.1. *A* natural *Lagrangian is a Lagrangian defined by*

$$L(\tau) = \frac{1}{2}K(\tau, \tau) - U \circ \pi(\tau),$$

where $U : Q \to \Re$ is the potential energy *function and $K \in T_2^0 Q$ is a* metric *on Q, which we will call the* kinetic inner product.

Definition 3.5.2. *The* kinetic energy *function* $K : TQ \to \mathbb{R}$ *for a natural Lagrangian is defined by*

$$K(\tau) = \frac{1}{2}K(\tau, \tau).$$

Here, K appearing with two arguments is the kinetic inner product. It will be clear from context which K is intended.

Lemma 3.5.1. *For a natural Lagrangian we have*

$$\langle FL(\tau), \rho\rangle = K(\tau, \rho), \ \forall \tau, \rho \in TQ.$$

Proof. We may calculate

$$\begin{aligned} \langle FL(\tau), \rho\rangle &= \left.\frac{d}{ds}\right|_{s=0} L(\tau + s\rho) \\ &= \left.\frac{d}{ds}\right|_{s=0} \left\{\frac{1}{2}K(\tau + s\rho, \tau + s\rho)\right\} \\ &= K(\tau, \rho). \end{aligned}$$

□

Lemma 3.5.2. *For a natural Lagrangian, $FL : TQ \to T^*Q$ is a diffeomorphism.*

Proof. This follows from the previous lemma, the positive definiteness of K and the finite dimension of Q. □

Lemma 3.5.3. *For a natural Lagrangian we have*

$$F^2L(\tau) = K, \ \forall \tau \in TQ.$$

Proof. We may calculate

$$\begin{aligned} F^2L(\tau)(\rho, \sigma) &= \left.\frac{\partial^2}{\partial u \partial s}\right|_{u=0, s=0} L(\tau + u\sigma + s\rho) \\ &= \left.\frac{\partial^2}{\partial u \partial s}\right|_{u=0, s=0} \left\{\frac{1}{2}K(\tau + u\sigma + s\rho, \tau + u\sigma + s\rho)\right\} \\ &= K(\rho, \sigma). \end{aligned}$$

□

The last lemma shows that L is indeed a Lagrangian.

Lemma 3.5.4. *For a natural Lagrangian L, we have*

$$E(\tau) = K(\tau) + U \circ \pi(\tau).$$

Proof. We simply calculate

$$\begin{aligned} E(\tau) &= \langle FL(\tau), \tau\rangle - L(\tau) \\ &= K(\tau, \tau) - \left\{\frac{1}{2}K(\tau, \tau) - U \circ \pi(\tau)\right\} \\ &= K(\tau) + U \circ \pi(\tau). \end{aligned}$$

□

3.6 Hamiltonian Systems

Definition 3.6.1. *A Hamiltonian system is a nonholonomic system for which $FL : TQ \to T^*Q$ is a diffeomorphism, the magnetic form M is closed ($dM = 0$ everywhere) and $D = TQ$ holds.*

Under these conditions, the intrinsic form of the equations of motion becomes

$$\ddot{q}(t) \lrcorner \left(d\theta_L + \frac{1}{2}\pi^* M\right) + dE = 0.$$

The 2-form $\omega = d\theta_L + \frac{1}{2}\pi^* M$ is a *symplectic form* for the manifold TQ. The energy function E on TQ is called the *Hamiltonian function* for the system.

Hamiltonian systems have been widely studied, and have many important properties, which we will not discuss here. We wish only to point out that it is very significant when a nonholonomic system can be reduced to a Hamiltonian system.

Chapter 4

Quasi-Velocities

We use the symbols and notation of chapter 3.

4.1 Momentum

The notion of *momentum* is a useful reference point in discussions of mechanical systems.

Definition 4.1.1. *Let α be a vector field defined on some region $W \subseteq Q$. The* momentum *associated with the vector α is the function $p^\alpha : TW \to \Re$ defined by*

$$p^\alpha(\tau) = \langle FL(\tau), \alpha \circ \pi(\tau) \rangle .$$

4.2 Quasi-Velocities

Suppose $\exists$ vector fields $\alpha_1, \ldots, \alpha_n$ defined on some region $W \subseteq Q$, n being the dimension of Q, such that the vector fields are linearly independent at each point of W.

At this point, we ask the reader to recall the definition of the *Lie bracket* of two vector fields. (see for example [33] or [6]). Let $C^k_{ij} : W \to \Re$ be the *structure functions* given by

$$[\alpha_i, \alpha_j] = C^k_{ij}\alpha_k.$$

Define functions $\mu^i : TW \to \Re$ by

$$\tau = \mu^i(\tau)\alpha_i, \forall \tau \in TW.$$

Observe that given local coordinates $q^i : W_0 \subseteq W \to \Re$, we have

$$\tau = \dot{q}^i(\tau)\frac{\partial}{\partial q^i}, \forall \tau \in TW_0.$$

Introducing the notation $\alpha_i^j = \dot{q}^j(\alpha_i)$, as is usually done, we have

$$\begin{aligned} \dot{q}^i(\tau) &= \dot{q}^i(\mu^j(\tau)\alpha_j) \\ &= \mu^j(\tau)\alpha_j^i. \end{aligned}$$

The μ^i are called *quasi-velocities*. We will call $\{\alpha_1, \ldots, \alpha_n\}$ a *quasi-velocity basis*.

Example. Body Angular Velocity

Recall the description of body angular velocity in section 2.5. The manifold there is $SO(3) \times \Re^3$, which we have parameterized by (W, x). If $(\dot{W}, \dot{x})$ is in the tangent space at (W, x) (the differentiation here would be with respect to the parameter of an appropriate curve), then the functions ω_1, ω_2, ω_3, $\dot{x}_1$, $\dot{x}_2$ and $\dot{x}_3$, determined by

$$\dot{W} = WA(\omega_1 e_1 + \omega_2 e_2 + \omega_3 e_3)$$

and

$$\dot{x} = \dot{x}_1 e_1 + \dot{x}_2 e_2 + \dot{x}_3 e_3,$$

are quasi-velocities. The basis vector fields are $(WA(e_1), 0)$, $(WA(e_2), 0)$, $(WA(e_3), 0)$, $(0, e_1)$, $(0, e_2)$ and $(0, e_3)$, respectively. These vector fields are globally defined. We will call this basis the *body frame basis*.

In order to calculate the structure functions, C_{ij}^k, we must calculate Lie brackets. In order to do this we must think of the basis vector fields as operators of differentiation. We will use the notation $\widehat{(WA(e_i), 0)}$ and $\widehat{(0, e_i)}$

to identify this interpretation. The most complex of these calculations is

$$
\begin{aligned}
&[\widehat{(W^* A(\epsilon_i), 0)}, \widehat{(W^* A(\epsilon_j), 0)}](W^*, x) \\
&\quad = \widehat{(W^* A(\epsilon_i), 0)}\{\widehat{(W^* A(\epsilon_j), 0)}(W^*, x)\} \\
&\qquad\qquad\qquad - \widehat{(W^* A(\epsilon_j), 0)}\{\widehat{(W^* A(\epsilon_i), 0)}(W^*, x)\} \\
&\quad = \widehat{(W^* A(\epsilon_i), 0)}(W^* A(\epsilon_j), 0) - \widehat{(W^* A(\epsilon_j), 0)}(W^* A(\epsilon_i), 0) \\
&\quad = (W^* A(\epsilon_i) A(\epsilon_j), 0) - (W^* A(\epsilon_j) A(\epsilon_i), 0) \\
&\quad = (W^* \{A(\epsilon_i) A(\epsilon_j) - A(\epsilon_j) A(\epsilon_i)\}, 0) \\
&\quad = (W^* A(A(\epsilon_i)\epsilon_j), 0) \\
&\quad = (W^* A(\varepsilon_{ij}^k \epsilon_k), 0) \\
&\quad = \varepsilon_{ij}^k \widehat{(W^* A(\epsilon_k), 0)}(W^*, x).
\end{aligned}
$$

where ε_{ij}^k is the three index *permutation symbol*, with its parity fixed by requiring $\varepsilon_{12}^3 = 1$. All other brackets vanish. □

Example. Spatial Angular Velocity

In the previous example, we could instead have determined functions ν_1, ν_2, ν_3, $\dot{x}_1$, $\dot{x}_2$ and $\dot{x}_3$ by

$$\dot{W}^* = A(\nu_1\epsilon_1 + \nu_2\epsilon_2 + \nu_3\epsilon_3)W^*$$

and

$$\dot{x} = \dot{x}_1\epsilon_1 + \dot{x}_2\epsilon_2 + \dot{x}_3\epsilon_3.$$

These again are quasi-velocities. The basis now includes $(A(\epsilon_1)W^*, 0)$, $(A(\epsilon_2)W^*, 0)$, $(A(\epsilon_3)W^*, 0)$, $(0, \epsilon_1)$, $(0, \epsilon_2)$ and $(0, \epsilon_3)$. We will call this basis the *spatial frame basis*. □

4.3 Poincaré's Equations

Now Lagrange's equations are

$$
\begin{aligned}
\left(\frac{\partial L}{\partial \dot{q}^i}\right)' - \frac{\partial L}{\partial q^i} - M_{ij}\dot{q}^j &= \lambda_j B_i^j \\
B_i^j \dot{q}^i &= b^j.
\end{aligned}
$$

where we think of L in this equation as being of the form $L = L(q, \dot{q})$.

We could also think of L as depending on q and μ. To this end we write

$$\begin{aligned} \overline{L} &= \overline{L}(q, \mu) \\ &= L(q, \dot{q}(q, \mu)) \\ &= L(q, \mu^j \alpha^i_j(q)). \end{aligned}$$

Then we obtain

$$\frac{\partial \overline{L}}{\partial q^i} = \frac{\partial L}{\partial q^i} + \frac{\partial L}{\partial \dot{q}^j} \mu^k \frac{\partial \alpha^j_k}{\partial q^i}$$

and

$$\frac{\partial \overline{L}}{\partial \mu^i} = \frac{\partial L}{\partial \dot{q}^j} \alpha^j_i.$$

and so

$$\begin{aligned} \left(\frac{\partial \overline{L}}{\partial \mu^i}\right)' &= \left(\frac{\partial L}{\partial \dot{q}^j}\right)' \alpha^j_i + \frac{\partial L}{\partial \dot{q}^j} \frac{\partial \alpha^j_i}{\partial q^k} \dot{q}^k \\ &= \left(\frac{\partial L}{\partial \dot{q}^j}\right)' \alpha^j_i + \frac{\partial L}{\partial \dot{q}^j} \frac{\partial \alpha^j_i}{\partial q^k} \mu^m \alpha^k_m. \end{aligned}$$

We may now calulate

$$\begin{aligned} &\lambda_j B^j_k \alpha^k_i + M_{km} \alpha^k_i \alpha^m_j \mu^j \\ &= \left\{\lambda_j B^j_k + M_{km} \dot{q}^m\right\} \alpha^k_i \\ &= \left\{\left(\frac{\partial L}{\partial \dot{q}^j}\right)' - \frac{\partial L}{\partial q^j}\right\} \alpha^j_i \\ &= \left(\frac{\partial \overline{L}}{\partial \mu^i}\right)' - \frac{\partial L}{\partial \dot{q}^j} \frac{\partial \alpha^j_i}{\partial q^k} \mu^m \alpha^k_m - \frac{\partial \overline{L}}{\partial q^j} \alpha^j_i + \frac{\partial L}{\partial \dot{q}^j} \mu^m \frac{\partial \alpha^j_m}{\partial q^k} \alpha^k_i \\ &= \left(\frac{\partial \overline{L}}{\partial \mu^i}\right)' + \frac{\partial L}{\partial \dot{q}^j} \mu^m [\alpha_i, \alpha_m]^j - \alpha^j_i \frac{\partial \overline{L}}{\partial q^j} \\ &= \left(\frac{\partial \overline{L}}{\partial \mu^i}\right)' + \frac{\partial L}{\partial \dot{q}^j} \mu^m C^k_{im} \alpha^j_k - \alpha^j_i \frac{\partial \overline{L}}{\partial q^j} \\ &= \left(\frac{\partial \overline{L}}{\partial \mu^i}\right)' + \frac{\partial \overline{L}}{\partial \mu^k} C^k_{ij} \mu^j - \alpha^j_i \frac{\partial \overline{L}}{\partial q^j}. \end{aligned}$$

If we write

$$\begin{aligned} \alpha_i \overline{L} &= \alpha_i^j \frac{\partial \overline{L}}{\partial q^j}, \\ \overline{B}_i^j &= \langle B^j, \alpha_i \rangle = B_k^j \alpha_i^k \qquad \text{and} \\ \overline{M}_{ij} &= M(\alpha_i, \alpha_j) = M_{km} \alpha_i^k \alpha_j^m, \end{aligned}$$

then on W_0, we have the equations

$$\left(\frac{\partial \overline{L}}{\partial \mu^i}\right)' + \frac{\partial \overline{L}}{\partial \mu^k} C_{ij}^k \mu^j - \overline{M}_{ij}\mu^j - \alpha_i \overline{L} = \lambda_j \overline{B}_i^j \tag{4.1}$$

$$\overline{B}_j^i \mu^j = b^i \tag{4.2}$$

These equations are called *Poincaré's equations.*

In light of the manner in which α_i^j transforms under coordinate transformations, $\alpha_i \overline{L}$ may be extended to all of W. We may also shrink W to the extent necessary for $B^1, \ldots, B^r$ to be extended to all of W. Thus $\overline{B}_i^j$ is defined on all of W.

So Poincaré's Equations apply to all of W, the intersection of the domains of α_i and B^j.

Note also that

$$\begin{aligned} \frac{\partial \overline{L}}{\partial \mu^i}(\tau) &= \left.\frac{d}{ds}\right|_{s=0} L(\tau + s\alpha_i \circ \pi(\tau)) \\ &= \langle FL(\tau), \alpha_i \circ \pi(\tau) \rangle \\ &= p^{\alpha_i}(\tau) \end{aligned} \tag{4.3}$$

holds (recall chapter 4.1, and the last expression is also defined on all of W.

Example. Euler's Equations

We will build on the preceding example. We consider a rigid body with no constraints or potentials present. The Lagrangian (equation 2.9 on page 17) is

$$L = \frac{1}{2} m \dot{x}^T \dot{x} + \frac{1}{2} \omega^T J \omega.$$

We have previously found the structure functions. Equations 4.1 and 4.2 become

$$\begin{aligned} J\dot{\omega} + A(\omega) J\omega &= 0 \qquad \text{and} \\ m\ddot{x} &= 0. \end{aligned}$$

The first of these constitutes *Euler's equations* (see for example [2]), which more typically are written

$$\begin{aligned} J_1\dot{\omega}_1 &= (J_2 - J_3)\omega_2\omega_3, \\ J_2\dot{\omega}_2 &= (J_3 - J_1)\omega_3\omega_1 \qquad \text{and} \\ J_3\dot{\omega}_3 &= (J_1 - J_2)\omega_1\omega_2. \end{aligned}$$

□

Example. Body Angular Momentum

We will again build on the rigid body/angular velocity example.

Equation 4.3 provides a convenient way to find the momenta associated with the basis vector fields. The momentum associated with $(WA(\epsilon_i), 0)$ is

$$\frac{\partial L}{\partial \omega^i} = \epsilon_i^T J\omega.$$

The vector $J\omega$ is usually called the *body angular momentum*, or *angular momentum in the body reference frame*. The momentum associated with $(0, \epsilon_i)$ is

$$\frac{\partial L}{\partial \dot{x}^i} = m\dot{x}^i.$$

The vector $m\dot{x}$ is the usual momentum for the center of mass. □

Example. Spatial Angular Momentum

We expand once more on the rigid body/angular velocity example.

In terms of the spatial frame basis, the Lagrangian (equation 2.9 on page 17) is

$$\begin{aligned} L &= \frac{1}{2}m\dot{x}^T\dot{x} + \frac{1}{2}\omega^T J\omega \\ &= \frac{1}{2}m\dot{x}^T\dot{x} + \frac{1}{2}\nu^T WJW^T\nu. \end{aligned}$$

The momentum associated with $(A(\epsilon_i)W, 0)$ is

$$\begin{aligned} \frac{\partial L}{\partial \nu^i} &= \epsilon_i^T WJW^T\nu \\ &= \epsilon_i^T WJ\omega. \end{aligned}$$

The vector $WJ\omega$ is usually called the *spatial angular momentum*, or *angular momentum in the spatial reference frame*. □

4.4 Natural Lagrangians

We suppose now that L is a natural Lagrangian, as in definition 3.5.1. We have

$$L(\tau) = \frac{1}{2}K(\tau,\tau) - U \circ \pi(\tau).$$

Then we have

$$\begin{aligned} \overline{L}(q,\mu) &= L(\mu^i \alpha_i(q)) \\ &= \frac{1}{2}K(\alpha_i(q),\alpha_j(q))\mu^i\mu^j - U(q). \end{aligned}$$

But now

$$\frac{\partial \overline{L}}{\partial \mu^k} = K(\alpha_k,\alpha_j)\mu^j,$$

so the left side of equation 4.1 is

$$\begin{aligned} L.S. &= \{K(\alpha_i,\alpha_j)\mu^j\}' + \{K(\alpha_k,\alpha_m)\mu^m\}C_{ij}^k\mu^j \\ &\quad -M(\alpha_i,\alpha_j)\mu^j - \frac{1}{2}\alpha_i\{K(\alpha_j,\alpha_k)\}\mu^j\mu^k + \alpha_i U \\ &= \{K(\alpha_i,\alpha_j)\mu^j\}' + K(\alpha_m,[\alpha_i,\alpha_j])\mu^m\mu^j - \frac{1}{2}\alpha_i\{K(\alpha_j,\alpha_k)\}\mu^j\mu^k \\ &\quad -M(\alpha_i,\alpha_j)\mu^j + \alpha_i U \\ &= \{K(\alpha_i,\alpha_j)\mu^j\}' \\ &\quad +\frac{1}{2}\{K([\alpha_i,\alpha_j],\alpha_k) + K(\alpha_j,[\alpha_i,\alpha_k]) - \alpha_i\{K(\alpha_j,\alpha_k)\}\}\mu^j\mu^k \\ &\quad -M(\alpha_i,\alpha_j)\mu^j + \alpha_i U. \end{aligned}$$

At this point, we ask the reader to recall the definition of the *Lie derivative* of a tensor with respect to a vector field, and the definition of the *exterior derivative* of a form (see for example [33] or [6]). Using well known formulae for the Lie derivative, we have

$$\begin{aligned} \alpha_i\{K(\alpha_j,\alpha_k)\} &= \mathcal{L}_{\alpha_i}\{K(\alpha_j,\alpha_k)\} \\ &= (\mathcal{L}_{\alpha_i}K)(\alpha_j,\alpha_k) + K(\mathcal{L}_{\alpha_i}\alpha_j,\alpha_k) + K(\alpha_j,\mathcal{L}_{\alpha_i}\alpha_k) \\ &= (\mathcal{L}_{\alpha_i}K)(\alpha_j,\alpha_k) + K([\alpha_i,\alpha_j],\alpha_k) + K(\alpha_j,[\alpha_i,\alpha_k]) \end{aligned}$$

or

$$K([\alpha_i,\alpha_j],\alpha_k) + K(\alpha_j,[\alpha_i,\alpha_k]) - \alpha_i\{K(\alpha_j,\alpha_k)\} = -(\mathcal{L}_{\alpha_i}K)(\alpha_j,\alpha_k).$$

Using this, the left side of equation 4.1 is

$$L.S. = \{K(\alpha_i,\alpha_j)\mu^j\}' - \frac{1}{2}(\mathcal{L}_{\alpha_i}K)(\alpha_j,\alpha_k)\mu^j\mu^k - M(\alpha_i,\alpha_j)\mu^j + \alpha_i U.$$

Poincaré's Equations may be summarized now either as

$$\begin{aligned} &\{K(\alpha_i,\alpha_j)\mu^j\}' \\ (4.4)\qquad &- \frac{1}{2}\{K([\alpha_i,\alpha_j],\alpha_k) + K(\alpha_j,[\alpha_i,\alpha_k]) - \alpha_i\{K(\alpha_j,\alpha_k)\}\}\mu^j\mu^k \\ &- M(\alpha_i,\alpha_j)\mu^j + \alpha_i U \\ &= \lambda_j \left\langle B^j,\alpha_i\right\rangle \qquad \text{and} \end{aligned}$$

$$(4.5)\qquad \left\langle B^i,\alpha_j\right\rangle \mu^j = b^i.$$

or as

$$(4.6)\qquad \{K(\alpha_i,\alpha_j)\mu^j\}' - \frac{1}{2}(\mathcal{L}_{\alpha_i}K)(\alpha_j,\alpha_k)\mu^j\mu^k - M(\alpha_i,\alpha_j)\mu^j + \alpha_i U = \lambda_j \left\langle B^j,\alpha_i\right\rangle \qquad \text{and}$$

$$(4.7)\qquad \left\langle B^i,\alpha_j\right\rangle \mu^j = b^i.$$

Chapter 5

Noether's Theorem

We again use the symbols and notation of chapter 3.

5.1 The General Result

First, observe that in coordinates,

$$\begin{aligned} FL &= \frac{\partial L}{\partial \dot{q}^i} dq^i \\ &= \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i \end{aligned}$$

holds. Then recall that Lagrange's equations are

$$\begin{aligned} \left(\frac{\partial L}{\partial \dot{q}^i}\right)' - \frac{\partial L}{\partial q^i} - M_{ij}\dot{q}^j &= \lambda_j B_i^j \\ B_i^j \dot{q}^i &= b^j. \end{aligned}$$

Let σ be a vector field on Q, with flow σ_s. Writing $\overline{q}(s, q) = \sigma_s(q)$, so that $\dot{\overline{q}} = T\sigma_s(\dot{q})$ holds, or in local coordinates

$$\dot{\overline{q}}^i(s, \dot{q}) = \frac{\partial \overline{q}^i}{\partial q^j}(s, q)\dot{q}^j(\dot{q}).$$

we may calculate

$$
\begin{aligned}
&\left(\frac{\partial L}{\partial \dot q^i}(\dot{\overline{q}})\frac{\partial \overline{q}^i}{\partial s}\right)' \\
&= \left(\frac{\partial L}{\partial \dot q^i}(\dot{\overline{q}})\right)'\frac{\partial \overline{q}^i}{\partial s} + \frac{\partial L}{\partial \dot q^i}(\dot{\overline{q}})\frac{\partial}{\partial q^j}\left(\frac{\partial \overline{q}^i}{\partial s}\right)\dot q^j \\
&= \left(\frac{\partial L}{\partial q^i}(\dot{\overline{q}}) + M_{ij}(\overline{q})\dot{\overline{q}}^j + \lambda_j B_i^j(\overline{q})\right)\frac{\partial \overline{q}^i}{\partial s} + \frac{\partial L}{\partial \dot q^i}(\dot{\overline{q}})\frac{\partial}{\partial s}\left(\frac{\partial \overline{q}^i}{\partial q^j}\dot q^j\right) \\
&= \frac{\partial}{\partial s}(\mathsf{L}(\dot{\overline{q}})) + M_{ij}(\overline{q})\frac{\partial \overline{q}^i}{\partial s}\dot{\overline{q}}^j + \lambda_j B_i^j(\overline{q})\frac{\partial \overline{q}^i}{\partial s}.
\end{aligned}
$$

Evaluating this at $s = 0$ we have

$$
\begin{aligned}
(p^\sigma(\dot q))' &= \langle FL(\dot q).\sigma(q)\rangle' \\
&= \left(\frac{\partial L}{\partial \dot q^i}(\dot q)\sigma^i(q)\right)' \\
&= \left.\frac{\partial}{\partial s}\right|_{s=0}(L\circ T\sigma_s(\dot q)) + M(\sigma(q).\dot q) + \lambda_j\left\langle B^j(q).\sigma(q)\right\rangle.
\end{aligned}
$$

where p^σ is the momentum associated with σ.

We need a couple of definitions before stating the result following from this calculation.

Definition 5.1.1. *We say that σ* fully preserves L *if*

$$\left.\frac{\partial}{\partial s}\right|_{s=0}(L\circ T\sigma_s(\tau)) = 0.\forall \tau \in TQ.$$

Definition 5.1.2. *We say that σ* partially preserves L *if*

$$\left.\frac{\partial}{\partial s}\right|_{s=0}(L\circ T\sigma_s(\tau)) = 0.\forall \tau \in D_\beta.$$

where, as before, $D_\beta = \{\tau \in TQ|\tau - \beta\circ\pi(\tau) \in D\}$.

We will only apply Noether's theorem to systems for which the magnetic field is identically zero.

Proposition 5.1.1. *For the $M \equiv 0$ case, if $\sigma \in D$ at each point, and σ partially preserves L, then the momentum $p^\sigma(\dot q) = \langle FL(\dot q).\sigma(q)\rangle$ associated with σ is a constant of the motion.*

The case in which $M \equiv dX$ holds with

$$\left.\frac{\partial}{\partial s}\right|_{s=0} X \circ T\sigma_s(\tau) = 0, \forall \tau \in D_\beta,$$

can also be shown to result in a constant. The constant in this case is $p^\sigma(\dot{q}) + X(\dot{q})$. When M is not exact, the situation is more complex, and we will not address it here.

5.2 Natural Lagrangians

We suppose now that L is a natural Lagrangian as in definition 3.5.1, so we have

$$L(\tau) = \frac{1}{2}K(\tau,\tau) - U \circ \pi(\tau).$$

Then we have

$$\begin{aligned} L \circ T\sigma_s(\tau) &= \frac{1}{2}K(T\sigma_s(\tau), T\sigma_s(\tau)) - U \circ \pi \circ T\sigma_s(\tau) \\ &= \frac{1}{2}K(T\sigma_s(\tau), T\sigma_s(\tau)) - U \circ \sigma_s \circ \pi(\tau). \end{aligned}$$

Recalling again the definition of the *Lie derivative* ([33] or [6]), this gives

$$\left.\frac{d}{ds}\right|_{s=0} (L \circ T\sigma_s(\tau)) = \frac{1}{2}(\mathfrak{L}_\sigma K)(\tau,\tau) - (\sigma U) \circ \pi(\tau).$$

So we obtain a constant of the motion, providing that we have

- $M \equiv 0$,
- $\sigma \in D$ and
- $\frac{1}{2}(\mathfrak{L}_\sigma K)(\tau,\tau) - (\sigma U) \circ \pi(\tau) = 0, \forall \tau \in D_\beta$.

Observe also, that for natural Lagrangians, the conserved momentum $p^\sigma(\tau) = \langle FL(\tau), \sigma \circ \pi(\tau)\rangle$ becomes $p^\sigma(\tau) = K(\tau, \sigma \circ \pi(\tau))$.

Chapter 6

Nonholonomic Systems with Symmetry

As usual, we use the symbols and notation of chapter 3.

6.1 Definition of Group Symmetry

Let G be a Lie group with a left action, $G \times Q \to Q$, on Q. Let $\Lambda : Q \to Q/G$ be the quotient projection.

We assume the Q/G has a manifold structure, except on a finite set P of isolated points (in the topological sense), and that local sections exist at all points in $Q - \Lambda^{-1}(P)$. We also assume that the action $G \times Q \to Q$ is free at all points in $Q - \Lambda^{-1}(P)$.

Definition 6.1.1. *We say that a vector field α is* group invariant *if and only if $\alpha \circ l_g = Tl_g \circ \alpha, \forall g \in G$ holds, where $l_g : Q \to Q$ is given by $q \mapsto gq$.*

Definition 6.1.2. *We say that G is a symmetry group for the nonholonomic system (Q, L, M, D, β) if and only if:*

1. *L is* group invariant. *That is to say that $L \circ Tl_g = L, \forall g \in G$ holds.*
2. *M is* group invariant. *That is to say that $M(Tl_g(\xi), Tl_g(\zeta)) = M(\xi, \zeta)$, $\forall g \in G, \xi \in TQ, \zeta \in TQ$ holds.*
3. *The constraint distribution D is group invariant. That is to say that $\tau \in D$ implies $Tl_g(\tau) \in D, \forall g \in G$.*

4. The vector field β is group invariant.

Recall that a natural Lagrangian is one of the form

$$L(\tau) = \frac{1}{2}K(\tau, \tau) - V \circ \pi(\tau).$$

Proposition 6.1.1. *A natural Lagrangian L is group invariant if and only if the kinetic inner product K and the potential energy V are each group invariant. That is to say if and only if*

- $K(Tl_g(\tau), Tl_g(\tau)) = K(\tau, \tau), \forall g \in G, \tau \in TQ$ *and*
- $V \circ l_g = V, \forall g \in G$

each holds.

Proof. We must have

$$\frac{1}{2}s^2 K(Tl_g(\tau), Tl_g(\tau)) + V \circ l_g \circ \pi(\tau) = \frac{1}{2}s^2 K(\tau, \tau)) + V \circ \pi(\tau), \forall g \in G, \tau \in TQ, s \in \Re.$$

But two polynomials are identical only if their coefficients are. □

Observe that in the presence of such a symmetry, D_β will also be group invariant, and that we have

$$\begin{aligned} \langle FL \circ Tl_g(\tau), Tl_g(\rho)\rangle &= \left.\frac{\partial}{\partial s}\right|_{s=0} L(Tl_g(\tau) + s(Tl_g(\rho)) \\ &= \left.\frac{\partial}{\partial s}\right|_{s=0} L(\tau + s\rho) \\ &= \langle FL(\tau), \rho\rangle, \forall g, \tau, \rho. \end{aligned}$$

Proposition 6.1.2. *The action of a symmetry group takes a trajectory of the nonholonomic system to another trajectory.*

Proof. This follows in an obvious way from Hamilton's Principle, proposition 3.4.1 on page 28, when one realises that, for a trajectory $q = q(t)$, $\delta(l_g \circ q) = Tl_g(\delta q)$ and $d(l_g \circ q)/dt = Tl_g(dq/dt)$ hold. □

So we see that we can, in principle, reduce the problem of finding trajectories to a problem on TQ/G rather than TQ. The trajectories in TQ/G will of course be restricted to lie in D_3/G.

Example. Nonholonomically Constrained Particle in $\mathbb{R}^3$

We introduce now an example which we will use to illustrate the ideas that follow.

This example has previously been used both in [5], and also in [7] to illustrate symmetry related theories for nonholonomic systems.

In this case Q is $\mathbb{R}^3$. We use (x, y, z) to parameterize Q. The Lagrangian is

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

The constraint is

$$\dot{z} = y\dot{x}.$$

The symmetry group G is E^2. We parameterize G by $(\Delta x, \Delta z)$. The group action is given by the map

$$(\overline{x}, \overline{y}, \overline{z}) = (x + \Delta x, y, z + \Delta z).$$

Differentiating in order to lift the action to TQ we have

$$(\dot{\overline{x}}, \dot{\overline{y}}, \dot{\overline{z}}) = (\dot{x}, \dot{y}, \dot{z}),$$

which shows that L and the constraint are preserved. □

6.2 Distributions Associated with Group Symmetry

Let $\mathfrak{g}$ be the Lie algebra of G. Recall the following well known definition.

Definition 6.2.1. *Let $\xi \in \mathfrak{g}$. The* fundamental vector field ξ_Q *on Q associated with ξ is defined by*

$$\xi_Q(q) = \left.\frac{d}{ds}\right|_{s=0} ((\exp s\xi)q), \forall q \in Q.$$

Definition 6.2.2. *The* vertical distribution V *on* Q *is given by*

$$V = \{\xi_Q(q) | \xi \in \mathfrak{g}, q \in Q\}.$$

This last definition is perhaps a contradiction in terms. Since the group action is not necessarily free everywhere on Q, the dimension of V may collapse at points in $\Lambda^{-1}(P)$. We will nevertheless call V a distribution.

Observe that V is group invariant, since $Tl_g(\xi_Q(q)) = (Ad_g\xi)_Q(gq), \forall g \in G, q \in Q$ holds.

Definition 6.2.3. *The distribution* S *on* Q *is given by* $S = V \cap D$.

Observe that the dimension of S may also collapse at points in $\Lambda^{-1}(P)$, and that S is also group invariant.

Recall that for natural Lagrangians, there is an associated *kinetic inner product* K. We may define the kinetic inner product K for a general Lagrangian L by $K(q) = F^2L(0_q)$, where 0_q is the zero vector at q.

Definition 6.2.4. *The* horizontal distribution H *on* Q *is given by* $H = D \cap S^{\perp}$, *where the kinetic inner product* K *is used to determine orthogonality.*

Observe that H is also group invariant, that $D = H \dotplus S$ holds, and that the dimension of H may jump at points in $\Lambda^{-1}(P)$.

Definition 6.2.5. *The distribution* N *on* Q *is given by* $N = V \cap S^{\perp}$, *where the kinetic inner product* K *is used to determine orthogonality.*

Observe that N is also group invariant, that $V = S \dotplus N$ holds, and that the dimension of N may jump or collapse at points in $\Lambda^{-1}(P)$.

Lemma 6.2.1. *The distributions* H, S *and* N *satisfy* $H \dotplus S \dotplus N = D + V$.

Proof. We have already observed that $D = H \dotplus S$ and $V = S \dotplus N$ hold. But we also have $H \cap N \subseteq D \cap V = S$, so that $H \cap N \subseteq H \cap S = \emptyset$ holds. □

Definition 6.2.6. *The distribution* R *on* Q *is given by* $R = (D + V)^{\perp}$, *where the kinetic inner product* K *is used to determine orthogonality.*

Observe that R is also group invariant, that $TQ = (D + V) \dotplus R$ holds, and that the dimension of R may jump on $\Lambda^{-1}(P)$.

Proposition 6.2.1. *If the set P of singular points in Q/G is empty, then the distribution $H \oplus R$ on Q determines a principal bundle connection on $\pi : Q \to Q/G$.*

Proof. It is clear that $H \oplus R$ is group invariant, and that $TQ = (H \oplus R) \oplus V$ holds. □

Even if P is not empty, we will have $T\pi(H \oplus R) = T(Q/G)$ elsewhere, and many of the concepts of principal bundles, such as horizontal lifts, may be used on $Q/G - P$.

Definition 6.2.7. *We say that the group symmetry satisfies the* dimension assumption *if $D + V = TQ$ (that is* $\dim R = 0$*) holds.*

If the dimension assumption is satisfied and the set P is empty, then H will be the horizontal distribution for a principal bundle connection. This is the rationale for the expression *horizontal distribution* for H.

Example. Nonholonomically Constrained Particle in $\mathbb{R}^3$

Recall that, for this example, we had

$$
\begin{aligned}
L &= \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2), \\
\dot{z} &= y\dot{x} \qquad \text{and} \\
(\overline{x}, \overline{y}, \overline{z}) &= (x + \Delta x, y, z + \Delta z).
\end{aligned}
$$

The distribution D is spanned by the vector fields $(\dot{x}, \dot{y}, \dot{z}) = (1, 0, y)$ and $(\dot{x}, \dot{y}, \dot{z}) = (0, 1, 0)$. The distribution V is spanned by $(\dot{x}, \dot{y}, \dot{z}) = (1, 0, 0)$ and $(\dot{x}, \dot{y}, \dot{z}) = (0, 0, 1)$. The distribution $S = V \cap D$ is spanned by $(\dot{x}, \dot{y}, \dot{z}) = (1, 0, y)$.

The kinetic inner product is given by

$$
K((\dot{x}, \dot{y}, \dot{z}), (\underline{\dot{x}}, \underline{\dot{y}}, \underline{\dot{z}})) = \dot{x}\underline{\dot{x}} + \dot{y}\underline{\dot{y}} + \dot{z}\underline{\dot{z}}.
$$

The distribution $H = D \cap S^{\perp}$ is thus spanned by $(\dot{x}, \dot{y}, \dot{z}) = (0, 1, 0)$. The distribution $N = V \cap S^{\perp}$ is spanned by $(\dot{x}, \dot{y}, \dot{z}) = (-y, 0, 1)$.

In this case the dimension assumption is satisfied. □

6.3 Local Bases Aligned with the Symmetry

For the next definition, we use the ideas introduced in chapter 4.

Definition 6.3.1. *A quasivelocity basis* $\{\kappa_1, \kappa_2, \ldots, \gamma_1, \gamma_2, \ldots, \chi_1, \chi_2, \ldots, \varsigma_1, \varsigma_2, \ldots\}$ *is said to be* aligned with the group symmetry *if and only if*

- *Each vector field in the basis is group invariant.*
- $\{\kappa_1, \kappa_2, \ldots\}$ *is a local basis for* H.
- $\{\gamma_1, \gamma_2, \ldots\}$ *is a local basis for* S.
- $\{\chi_1, \chi_2, \ldots\}$ *is a local basis for* X.
- $\{\varsigma_1, \varsigma_2, \ldots\}$ *is a local basis for* R.

We may on occasion introduce such a basis without distinguishing the χ vector fields from the ς vector fields.

Proposition 6.3.1. *There is a local basis aligned with the symmetry at each* $q_0 \in Q - \Lambda^{-1}(P)$.

Proof. Let $\Sigma : W \subseteq Q/G \rightarrow Q$ be a local section of $\Lambda : Q \rightarrow Q/G$ at q_0. So we have $\Lambda \circ \Sigma = id$ and $\Sigma \circ \Lambda(q_0) = q_0$. Let $\Gamma : \Lambda^{-1}(W) \rightarrow G$ be the map such that $\Gamma(q)(\Sigma \circ \Lambda(q)) = q. \forall q \in \Lambda^{-1}(W)$ holds. By this we mean that $(\Lambda, \Gamma) : \Lambda^{-1}(W) \rightarrow W \times G$ gives the local trivialisation.

Observe now that any vector field $\alpha : W_0 \subseteq \Lambda^{-1}(W) \rightarrow TQ$ may be used to define a group invariant vector field $\overline{\alpha} : \Lambda^{-1}(\Sigma^{-1}(W_0)) \rightarrow TQ$ by

$$\overline{\alpha}(q) = Tl_{\Gamma(q)}(\alpha(\Sigma \circ \Lambda(q)))$$

The linear independence of a set of vector fields will be preserved since $Tl_{\Gamma(q)}$ is non-singular for $q \in \Lambda^{-1}(W) \subseteq Q - \Lambda^{-1}(P)$.

So the proposition is proved by taking any local basis for each of H, S, X and R, redefining the vector fields in the above manner, and then taking the union of these modified bases. □

Example. Nonholonomically Constrained Particle in $\mathbb{R}^3$

Recall that, for this example, the kinetic inner product is given by

$$K((\dot{x},\dot{y},\dot{z}),(\underline{\dot{x}},\underline{\dot{y}},\underline{\dot{z}})) = \dot{x}\underline{\dot{x}} + \dot{y}\underline{\dot{y}} + \dot{z}\underline{\dot{z}}.$$

with

$$\begin{aligned} H &= \operatorname{span}\{(0,1,0)\}, \\ S &= \operatorname{span}\{(1,0,y)\}, \\ N &= \operatorname{span}\{(-y,0,1)\} \qquad \text{and} \\ \dim R &= 0. \end{aligned}$$

So $\{(0,1,0),(1,0,y),(-y,0,1)\}$ is a basis aligned with the symmetry. □

Chapter 7

Finding Group Invariant Constants

7.1 Restrictions on the Vector Field

We have given in proposition 5.1.1 on page 40 a criterion for a constant of the motion of the form

$$C(\tau) = \langle FL(\tau).\sigma \circ \pi(\tau)\rangle$$

to exist.

A constant of this sort is of greater value in a problem with a symmetry group if it is group invariant - that is to say if $C \circ Tl_g = C.\forall g \in G$ holds. In that case it induces a function $\overline{C}$ on TQ/G which is a constant of the reduced motion.

We assume now that we have a symmetry group as in the previous section. We find then, if σ is itself group invariant,

$$\begin{aligned} C \circ Tl_g(\tau) &= \langle FL \circ Tl_g(\tau).\sigma \circ \pi \circ Tl_g(\tau)\rangle \\ &= \langle FL \circ Tl_g(\tau).\sigma \circ l_g \circ \pi(\tau)\rangle \\ &= \langle FL \circ Tl_g(\tau).Tl_g \circ \sigma \circ \pi(\tau)\rangle \\ &= \langle FL(\tau).\sigma \circ \pi(\tau)\rangle \\ &= C(\tau). \end{aligned} \tag{7.1}$$

Hence we will impose the additional requirement that σ be group invariant.

Hereafter in this section we assume that L is a natural Lagrangian and that there is no magnetic field (that is, $M \equiv 0$ holds) and that $\beta \equiv 0$ holds.

The requirements for σ now are (section 5.2):

- σ is group invariant.
- $\sigma \in D$.
- $\frac{1}{2}(\mathcal{L}_\sigma K)(\tau, \tau) = (\sigma U) \circ \pi(\tau), \forall \tau \in D$.

The constant of the motion now is $K(\tau, \sigma \circ \pi(\tau))$.

Since the potential U is group invariant, we will have $\sigma(q)U = 0, \forall q$ if σ is vertical. Hence we now also impose the requirement $\sigma \in S$, leaving us with

$$(\mathcal{L}_\sigma K)(\tau, \tau) = 0, \forall \tau \in D.$$

In order to examine this criterion, let $\{\kappa_1, \kappa_2, \ldots, \gamma_1, \gamma_2, \ldots, \lambda_1, \lambda_2, \ldots, \varsigma_1, \varsigma_2, \ldots\}$ be a basis aligned with the symmetry. The criterion above is equivalent to

$$(\mathcal{L}_\sigma K)(\alpha, \rho) \equiv 0, \ \forall \alpha, \rho \in \{\kappa_1, \kappa_2, \ldots, \gamma_1, \gamma_2, \ldots\}.$$

We may calculate

$$(\mathcal{L}_\sigma K)(\alpha, \rho) = \sigma(K(\alpha, \rho)) - (K([\sigma, \alpha], \rho) + K(\alpha, [\sigma, \rho])).$$

But $\sigma(K(\alpha, \rho))$ is identically zero, since σ is vertical and $K(\alpha, \rho)$ is a group invariant function on Q. So we need

$$K([\sigma, \alpha], \rho) + K(\alpha, [\sigma, \rho]) = 0, \ \forall \alpha, \rho \in \{\kappa_1, \kappa_2, \ldots, \gamma_1, \gamma_2, \ldots\}.$$

7.2 The Flatness Conditions

Since σ is vertical, we may write $\sigma = \sigma^i \gamma_i$. The functions σ^i will be group invariant since σ is. Making particular choices for α and ρ we require

$$\begin{aligned}
0 &= K([\sigma, \gamma_j], \gamma_k) + K(\gamma_j, [\sigma, \gamma_k]) \\
&= -(\gamma_j \sigma^i) K(\gamma_i, \gamma_k) + \sigma^i K([\gamma_i, \gamma_j], \gamma_k) \\
&\quad -(\gamma_k \sigma^i) K(\gamma_j, \gamma_i) + \sigma^i K(\gamma_j, [\gamma_i, \gamma_k]) \\
&= \sigma^i (K([\gamma_i, \gamma_j], \gamma_k) + K(\gamma_j, [\gamma_i, \gamma_k]))
\end{aligned}$$

and

$$\begin{aligned} 0 &= K([\sigma,\kappa_j],\kappa_k) + K(\kappa_j,[\sigma,\kappa_k]) \\ &= -(\kappa_j\sigma^i)K(\gamma_i,\kappa_k) + \sigma^i K([\gamma_i,\kappa_j],\kappa_k) \\ &\quad -(\kappa_k\sigma^i)K(\kappa_j,\gamma_i) + \sigma^i K(\kappa_j,[\gamma_i,\kappa_k]) \\ &= \sigma^i(K([\gamma_i,\kappa_j],\kappa_k) + K(\kappa_j,[\gamma_i,\kappa_k])). \end{aligned}$$

If we hope to find a non-zero vector field σ, we are led to the conditions:

$$K([\gamma_i,\gamma_j],\gamma_k) + K(\gamma_j,[\gamma_i,\gamma_k]) = 0 \tag{7.2}$$

$$K([\gamma_i,\kappa_j],\kappa_k) + K(\kappa_j,[\gamma_i,\kappa_k]) = 0. \tag{7.3}$$

Proposition 7.2.1. *If equations 7.2 and 7.3 are satisfied for one choice of aligned basis, they will be satisfied for any other.*

Proof. Suppose $\{\bar\kappa_1,\bar\kappa_2,\dots\}$ is another basis for H, and $\{\bar\gamma_1,\bar\gamma_2,\dots\}$ is another basis for S. We must have $\bar\kappa_i = a_i^j\kappa_j$ and $\bar\gamma_i = b_i^j\gamma_j$ where a_i^j and b_i^j are group invariant functions.

So we have

$$\begin{aligned} [\bar\gamma_i,\bar\gamma_j] &= b_i^k b_j^m[\gamma_k,\gamma_m] + b_i^k(\gamma_k b_j^m)\gamma_m - b_j^m(\gamma_m b_i^k)\gamma_k \\ &= b_i^k b_j^m[\gamma_k,\gamma_m] \end{aligned}$$

and then

$$\begin{aligned} K([\bar\gamma_i,\bar\gamma_j],\bar\gamma_k) &+ K(\bar\gamma_j,[\bar\gamma_i,\bar\gamma_k]) \\ &= b_i^m b_j^n b_k^p(K([\gamma_m,\gamma_n],\gamma_p) + K(\gamma_n,[\gamma_m,\gamma_p])) \\ &= 0. \end{aligned}$$

So equation 7.2 holds for the new basis.

Next we have

$$\begin{aligned} [\bar\gamma_i,\bar\kappa_j] &= b_i^k a_j^m[\gamma_k,\kappa_m] + b_i^k(\gamma_k a_j^m)\kappa_m - a_j^m(\kappa_m b_i^k)\gamma_k \\ &= b_i^k a_j^m[\gamma_k,\gamma_m] - a_j^m(\kappa_m b_i^k)\gamma_k \end{aligned}$$

and then

$$\begin{aligned} K([\bar\gamma_i,\bar\kappa_j],\bar\kappa_k) &+ K(\bar\kappa_j,[\bar\gamma_i,\bar\kappa_k]) \\ &= b_i^m a_j^n a_k^p(K([\gamma_m,\kappa_n],\kappa_p) + K(\kappa_n,[\gamma_m,\kappa_p])) \\ &= 0. \end{aligned}$$

since $K(\kappa_r,\gamma_s) = 0,\ \forall r,s$ holds. So equation 7.3 holds for the new basis. □

Definition 7.2.1. *We call equations 7.2 and 7.3 the* flatness conditions. *We will say that a symmetry satisfying the flatness conditions is* flat.

We will later describe physical systems of interest for which these conditions are satisfied.

Finally, observe that we have

$$\begin{aligned} 0 &= \gamma_i(K(\gamma_j,\gamma_k)) \\ &= (\mathcal{L}_{\gamma_i}K)(\gamma_j,\gamma_k) + K([\gamma_i,\gamma_j],\gamma_k) + K(\gamma_j,[\gamma_i,\gamma_k]) \end{aligned}$$

and similarly

$$\begin{aligned} 0 &= \gamma_i(K(\kappa_j,\kappa_k)) \\ &= (\mathcal{L}_{\gamma_i}K)(\kappa_j,\kappa_k) + K([\gamma_i,\kappa_j],\kappa_k) + K(\gamma_j,[\kappa_i,\kappa_k]), \end{aligned}$$

so that the flatness conditions, equations 7.2 and 7.3 may be restated as

$$(\mathcal{L}_{\gamma_i}K)(\gamma_j,\gamma_k) = 0 \tag{7.4}$$

$$(\mathcal{L}_{\gamma_i}K)(\kappa_j,\kappa_k) = 0 \tag{7.5}$$

Example. Nonholonomically Constrained Particle in $\mathbb{R}^3$

Recall that, for this example, the kinetic inner product is given by

$$K((\dot{x},\dot{y},\dot{z}),(\underline{\dot{x}},\underline{\dot{y}},\underline{\dot{z}})) = \dot{x}\underline{\dot{x}} + \dot{y}\underline{\dot{y}} + \dot{z}\underline{\dot{z}},$$

with

$$\begin{aligned} \kappa_1 &= (0,1,0) \qquad \text{and} \\ \gamma_1 &= (1,0,y). \end{aligned}$$

Equation 7.2 is true since S has dimension 1. For equation 7.3 we first calculate

$$\begin{aligned} [\gamma_1,\kappa_1] &= \gamma_1(0,1,0) - \kappa_1(1,0,y) \\ &= (0,0,0) - (0,0,1) \\ &= (0,0,-1), \end{aligned}$$

and then

$$K([\gamma_1,\kappa_1],\kappa_1) + K(\kappa_1,[\gamma_1,\kappa_1]) = 0.$$

So the flatness condition is satisfied. □

7.3 The PDE

Making the remaining choice for α and ρ above, we also require

$$K([\sigma,\gamma_j],\kappa_k) + K(\gamma_j,[\sigma,\kappa_k]) = 0. \tag{7.6}$$

Proposition 7.3.1. *If equation 7.6 is satisfied for one choice of aligned basis, it will be true for any other.*

Proof. Suppose $\{\overline{\kappa}_1,\overline{\kappa}_2,\dots\}$ is another basis for H, and $\{\overline{\gamma}_1,\overline{\gamma}_2,\dots\}$ is another basis for S. We must have $\overline{\kappa}_i = a_i^j\kappa_j$ and $\overline{\gamma}_i = b_i^j\gamma_j$ where a_i^j and b_i^j are group invariant functions. So we have

$$\begin{aligned}
K([\sigma,\overline{\gamma}_j],\overline{\kappa}_k) &+ K(\overline{\gamma}_j,[\sigma,\overline{\kappa}_k]) \\
&= K((\sigma b_j^m)\gamma_m + b_j^m[\sigma,\gamma_m], a_k^n\kappa_n) \\
&\quad + K(b_j^m\gamma_m, (\sigma a_k^n)\kappa_n + a_k^n[\sigma,\kappa_n]) \\
&= a_k^n b_j^m(K([\sigma,\gamma_m],\kappa_n) + K(\gamma_m,[\sigma,\kappa_n])) \\
&= 0.
\end{aligned}$$

So equation 7.6 holds for the new basis. ☐

Writing $\sigma = \sigma^i\gamma_i$ as before, with the functions σ^i group invariant, we have

$$\begin{aligned}
0 &= K([\sigma,\gamma_j],\kappa_k) + K(\gamma_j,[\sigma,\kappa_k]) \\
&= -(\gamma_j\sigma^i)K(\gamma_i,\kappa_k) + \sigma^i K([\gamma_i,\gamma_j],\kappa_k) \\
&\quad -(\kappa_k\sigma^i)K(\gamma_j,\gamma_i) + \sigma^i K(\gamma_j,[\gamma_i,\kappa_k]) \\
&= -K(\gamma_j,\gamma_i)(\kappa_k\sigma^i) + \sigma^i(K([\gamma_i,\gamma_j],\kappa_k) + K(\gamma_j,[\gamma_i,\kappa_k])).
\end{aligned}$$

or

$$K(\gamma_j,\gamma_i)(\kappa_k\sigma^i) = \sigma^i(K([\gamma_i,\gamma_j],\kappa_k) + K(\gamma_j,[\gamma_i,\kappa_k])). \tag{7.7}$$

Observe that in this last equation, the matrix with $K(\gamma_j,\gamma_i)$ for coefficients is positive definite, since K is, and therefore non-singular.

Also in equation 7.7, $K(\gamma_j,\gamma_i)$ and $K([\gamma_i,\gamma_j],\kappa_k)+K(\gamma_j,[\gamma_i,\kappa_k])$ are each group invariant functions on Q, since the Lie bracket of group invariant vector fields is group invariant, and K is group invariant. But $\kappa_k\sigma^i$ is also group invariant, as follows from

$$\begin{aligned}
\kappa_k(gq)\sigma^i &= (Tl_g\kappa_k(q))\sigma^i \\
&= \kappa_k(q)(\sigma^i \circ l_g) \\
&= \kappa_k(q)\sigma^i.
\end{aligned}$$

In fact, since κ_k is group invariant, it induces a vector field $\overline{\kappa}_k$ on Q/G such that $T\Lambda \circ \kappa_k = \overline{\kappa}_k \circ \Lambda$ holds. Similarly, σ^i induces a function $\overline{\sigma}^i$ on Q such that $\sigma^i = \overline{\sigma}^i \circ \Lambda$ holds. Moreover, the function induced on Q by $\kappa_k\sigma^i$ is precisely $\overline{\kappa}_k\overline{\sigma}^i$.

So we see that equation 7.7 is locally a partial differential equation on Q/G.

We will not discuss the integrability of this PDE in the general case.

Definition 7.3.1. *We say that a nonholonomic system is* elementary, *or that the system falls into the* elementary case, *if:*

- *The flatness conditions are satisfied.*
- *The distribution H is of dimension 1.*

In the elementary case, equation 7.7 is locally a linear ordinary differential equation. This O.D.E. will have parameters if Q/G has dimension greater than 1. In light of proposition 7.3.1, solutions may be extended beyond a local patch, to all of Q/G, perhaps excluding boundary points. There is also the possibility of a multiple valued solution. So we have:

Proposition 7.3.2. *In the elementary case, we may obtain linearly independent solutions of equation 7.6, equal in number to the dimension of S.*

We will later describe physical systems of considerable interest which fall into the elementary case.

As in the previous section, from

$$\begin{aligned} 0 &= \gamma_i(K(\gamma_j, \kappa_k)) \\ &= (\mathcal{L}_{\gamma_i} K)(\gamma_j, \kappa_k) + K([\gamma_i, \gamma_j], \kappa_k) + K(\gamma_j, [\gamma_i, \kappa_k]), \end{aligned}$$

we may restate equation 7.7 as

$$K(\gamma_j, \gamma_i)(\kappa_k\sigma^i) + \sigma^i(\mathcal{L}_{\gamma_i} K)(\gamma_j, \kappa_k) = 0. \tag{7.8}$$

Example. Nonholonomically Constrained Particle in $\Re^3$

Again, the kinetic inner product is given by

$$K((\dot{x}, \dot{y}, \dot{z}), (\underline{\dot{x}}, \underline{\dot{y}}, \underline{\dot{z}})) = \dot{x}\underline{\dot{x}} + \dot{y}\underline{\dot{y}} + \dot{z}\underline{\dot{z}}.$$

with

$$\begin{aligned} \kappa_1 &= (0, 1, 0) \qquad \text{and} \\ \gamma_1 &= (1, 0, y). \end{aligned}$$

We also had

$$[\gamma_1, \kappa_1] = (0, 0, -1).$$

For equation 7.7 we calculate

$$K(\gamma_1, \gamma_1) = 1 + y^2,$$

and then

$$K(\gamma_1, [\gamma_1, \kappa_1]) = -y,$$

in order to obtain the PDE

$$(1 + y^2)\frac{d\sigma^1}{dy} = \sigma^1(-y).$$

This equation is solved by

$$\sigma^1 = \frac{1}{\sqrt{1 + y^2}}.$$

The associated constant of the motion is

$$K((\dot{x}, \dot{y}, \dot{z}), \sigma^1\gamma_1) = \frac{\dot{x} + y\dot{z}}{\sqrt{1 + y^2}}.$$

□

7.4 Vertical Component of Trajectory

Suppose we have independent solutions of equation 7.6, equal in number to $\dim S$, for a system with a flat symmetry.

Then the equation

$$K(\Omega, \sigma_i(q)) = J_i$$

together with $\Omega \in S$, determines Ω for each value of q and J (thinking of J as a vector). If we write

$$\Omega = \Omega(J, q),$$

then from

$$\begin{aligned}
K(Tl_g(\Omega(J,q)), \sigma_i(gq)) &= K(Tl_g(\Omega(J,q)), Tl_g \circ \sigma_i(q)) \\
&= K(\Omega(J,q), \sigma_i(q)) \\
&= J_i \\
&= K(\Omega(J,gq), \sigma_i(gq)),
\end{aligned}$$

we see that

$$\Omega(J, gq) = Tl_g \circ \Omega(J, q)$$

holds.

If $q = q(t)$ is the trajectory of the system, then if we put

$$J_i(t) = K(\dot{q}(t), \sigma_i \circ q(t)),$$

J_i will be a constant function of t. If we split $\dot{q}$ into its H and S components, as

$$\dot{q} = \dot{q}_H + \dot{q}_S$$

with $\dot{q}_H \in H$ and $\dot{q}_S \in S$, we must therefore have

$$\dot{q}_S(t) = \Omega(J, q(t)).$$

7.5 Adjoint Equation

We now show that Ω, defined in section 7.4, satisfies a PDE, which may be described as being *adjoint* to equation 7.8. This PDE will be significant in the next section.

Let $\{\kappa_1, \kappa_2, \ldots, \gamma_1, \gamma_2, \ldots\}$ be an aligned basis for D, as usual. Writing $\sigma_i = \sigma_i^j \gamma_j$ and $\Omega = \Omega^k \gamma_k$, we have, using equation 7.8,

$$\begin{aligned}
0 &= \kappa_m(K(\Omega, \sigma_i)) \\
&= \{\kappa_m(K(\Omega, \gamma_j))\}\sigma_i^j + \Omega^k K(\gamma_k, \gamma_j)\{\kappa_m \sigma_i^j\} \\
&= \{\kappa_m(K(\Omega, \gamma_j))\}\sigma_i^j - \Omega^k \{\sigma_i^j (\mathcal{L}_{\gamma_j} K)(\gamma_k, \kappa_m)\} \\
&= \sigma_i^j \{\kappa_m(K(\gamma_j, \Omega)) - (\mathcal{L}_{\gamma_j} K)(\kappa_m, \Omega)\}.
\end{aligned}$$

We may conclude that Ω satisfies

$$\kappa_j(K(\gamma_i,\Omega)) = (\mathcal{L}_{\gamma_i}K)(\kappa_j,\Omega). \tag{7.9}$$

In light of

$$\begin{aligned} 0 &= \gamma_i(K(\kappa_j,\Omega)) \\ &= (\mathcal{L}_{\gamma_i}K)(\kappa_j,\Omega) + K([\gamma_i,\kappa_j],\Omega) + K(\kappa_j,[\gamma_i,\Omega]), \end{aligned}$$

equation 7.9 may be restated as

$$\kappa_j(K(\gamma_i,\Omega)) + K([\gamma_i,\kappa_j],\Omega) + K(\kappa_j,[\gamma_i,\Omega]) = 0. \tag{7.10}$$

Definition 7.5.1. *We call equation 7.9(or equation 7.10) the* adjoint equation *to equation 7.8(or equation 7.6, respectively).*

One other useful form of the adjoint equation is

$$\kappa_j(K(\gamma_i,\gamma_k)\Omega^k) + \{K([\gamma_i,\kappa_j],\gamma_k) + K(\kappa_j,[\gamma_i,\gamma_k])\}\Omega^k = 0. \tag{7.11}$$

Proposition 7.5.1. *If equation 7.10 is satisfied for one choice of aligned basis, it will be true for any other.*

Proof. Suppose $\{\overline{\kappa}_1,\overline{\kappa}_2,\dots\}$ is another basis for H, and $\{\overline{\gamma}_1,\overline{\gamma}_2,\dots\}$ is another basis for S. We must have $\overline{\kappa}_i = a_i^j\kappa_j$ and $\overline{\gamma}_i = b_i^j\gamma_j$ where a_i^j and b_i^j are group invariant functions. So we have

$$\begin{aligned} \overline{\kappa}_j(K(\overline{\gamma}_i,\Omega)) &= a_j^n\kappa_n(b_i^m K(\gamma_m,\Omega)) \\ &= a_j^n(\kappa_n b_i^m)K(\gamma_m,\Omega) + a_j^n b_i^m\kappa_n(K(\gamma_m,\Omega)) \end{aligned}$$

and at the same time

$$\begin{aligned} K([\overline{\gamma}_i,\overline{\kappa}_j],\Omega) &+ K(\overline{\kappa}_j,[\overline{\gamma}_i,\Omega]) \\ &= K(a_j^n b_i^m[\gamma_m,\kappa_n] - a_j^n(\kappa_n b_i^m)\gamma_m,\Omega) \\ &\quad + K(a_j^n\kappa_n, b_i^m[\gamma_m,\Omega] - (\Omega b_i^m)\gamma_m) \\ &= -a_j^n(\kappa_n b_i^m)K(\gamma_m,\Omega) \\ &\quad + a_j^n b_i^m\{K([\gamma_m,\kappa_n],\Omega) + K(\kappa_n,[\gamma_m,\Omega])\}. \end{aligned}$$

Adding these results we have

$$\begin{aligned} \overline{\kappa}_j(K(\overline{\gamma}_i,\Omega)) &+ K([\overline{\gamma}_i,\overline{\kappa}_j],\Omega) + K(\overline{\kappa}_j,[\overline{\gamma}_i,\Omega]) \\ &= a_j^n b_i^m\{\kappa_n(K(\gamma_m,\Omega)) + K([\gamma_m,\kappa_n],\Omega) + K(\kappa_n,[\gamma_m,\Omega])\} \\ &= 0. \end{aligned}$$

$\square$

The discussion of equation 7.7 in section 7.3 may be adapted to see that equation 7.10 is locally a partial differential equation on Q/G. As in section 7.3, in the elementary case, we may obtain independent solutions of equation 7.10, equal in number to the dimension of S.

Given these independent solutions to equation 7.10, we may reverse the process used to construct Ω, by using the equation

$$K(\Omega_i(q), \sigma) = J_i$$

together with $\sigma \in S$, to construct solutions to equation 7.6.

Example. Nonholonomically Constrained Particle in $\mathbb{R}^3$

Again, the kinetic inner product is given by

$$K((\dot{x}, \dot{y}, \dot{z}), (\underline{\dot{x}}, \underline{\dot{y}}, \underline{\dot{z}})) = \dot{x}\underline{\dot{x}} + \dot{y}\underline{\dot{y}} + \dot{z}\underline{\dot{z}},$$

with

$$\begin{aligned} \kappa_1 &= (0, 1, 0) \qquad \text{and} \\ \gamma_1 &= (1, 0, y). \end{aligned}$$

We also had

$$\begin{aligned} [\gamma_1, \kappa_1] &= (0, 0, -1), \\ K(\gamma_1, \gamma_1) &= 1 + y^2 \qquad \text{and} \\ K(\gamma_1, [\gamma_1, \kappa_1]) &= -y. \end{aligned}$$

For equation 7.11 we obtain

$$\frac{d}{dy}\left\{(1+y^2)\Omega^1\right\} = y\Omega^1.$$

This equation is solved by

$$\Omega^1 = \frac{C}{\sqrt{1+y^2}},$$

where C is a constant of integration. Thus we have

$$\begin{aligned} \Omega &= \Omega^1\gamma_1 \\ &= \frac{C}{\sqrt{1+y^2}}(1, 0, y). \end{aligned}$$

Since κ_1 has no $\dot{x}$ or $\dot{z}$ component, throughout the motion we have

$$\begin{aligned} \dot{x} &= \frac{C}{\sqrt{1+y^2}} \qquad \text{and} \\ \dot{z} &= \frac{Cy}{\sqrt{1+y^2}}. \end{aligned}$$

A simple calculation then gives

$$\frac{\dot{x}+y\dot{z}}{\sqrt{1+y^2}} = C,$$

so that C is the constant of the motion found earlier. □

Chapter 8

Equations of Motion

In this section we examine the equations of motion (that is Poincaré's equations) in the presence of a symmetry group, using a basis aligned with the symmetry, **for the case where L is a natural Lagrangian with there is no magnetic field (that is, $M \equiv 0$ holds) and $\beta \equiv 0$ holds.**

8.1 Reduction to TQ/G

Recall Poincaré's equations for natural Lagrangians, equations 4.6 and 4.7 on page 38,

$$
\begin{aligned}
&\{K(\alpha_i, \alpha_j)\mu^j\}' - \frac{1}{2}(\mathcal{L}_{\alpha_i} K)(\alpha_j, \alpha_k)\mu^j\mu^k \\
&\qquad\qquad - M(\alpha_i, \alpha_j)\mu^j + \alpha_i U = \lambda_j \langle B^j, \alpha_i \rangle \qquad \text{and} \\
&\langle B^i, \alpha_j \rangle \mu^j = b^i.
\end{aligned}
$$

If we now take the basis $\{\alpha_1, \alpha_2, \dots\}$ instead to be a basis $\{\zeta_1, \zeta_2, \dots, \chi_1, \chi_2, \dots\}$, such that $\{\zeta_1, \zeta_2, \dots\}$ is a basis for D, and correspondingly rename the quasi-velocities μ_j according to

$$\dot{q}^i(\tau) = \theta^j(\tau)\zeta_j^i + \varpi^j(\tau)\chi_j^i,$$

then the constraint equation reduces to

$$\langle B^i, \chi_j \rangle \{\varpi^j(\tau) - \varpi^j(\beta)\} = 0.$$

Now the matrix with coefficients given by $\langle B^i, \chi_j \rangle$ is non-singular, since $D \cap (\text{span}\{\chi_1, \chi_2, \dots\}) = \emptyset$ holds. So ϖ^j is determined (for each j) in the general case, and in this case we must have $\varpi^j \equiv 0$. Using this the remaining equation above becomes

$$\{K(\zeta_i, \zeta_j)\theta^j\}' - \frac{1}{2}(\mathcal{L}_{\zeta_i} K)(\zeta_j, \zeta_k)\theta^j\theta^k - M(\zeta_i, \zeta_j)\theta^j + \zeta_i U = 0. \tag{8.1}$$

If in place of equations 4.6 and 4.7, we use Poincaré's equations in the form of 4.4 and 4.5, this last equation becomes instead

$$\begin{aligned} &\{K(\zeta_i, \zeta_j)\theta^j\}' \\ &\quad - \frac{1}{2}\{K([\zeta_i, \zeta_j], \zeta_k) + K(\zeta_j, [\zeta_i, \zeta_k]) - \zeta_i\{K(\zeta_j, \zeta_k)\}\}\theta^j\theta^k \\ &\quad - M(\zeta_i, \zeta_j)\theta^j + \zeta_i U = 0 \end{aligned} \tag{8.2}$$

We will later use the following:

Lemma 8.1.1. *The collection of systems of equations, obtained from equation 8.1 (or from equation 8.2) by taking each local basis $\{\zeta_1, \zeta_2, \dots\}$ of D, determines the same trajectory in Q as the collection of systems of Lagrange's equations in definition 3.1.4 on page 20.*

Proof. We will only sketch the proof.

Introduce arbitrary coordinates q^i on Q. We must obtain equation 3.1 and equation 3.2. We immediately have

$$\begin{aligned} B^j_i \dot{q}^i &= \left\langle B^j, \dot{q}^i \frac{\partial}{\partial q^i} \right\rangle \\ &= \left\langle B^j, \theta^i \zeta_i \right\rangle \\ &= \theta^k \left\langle B^j, \zeta_i \right\rangle \\ &= 0, \end{aligned}$$

which is equation 3.2.

We now write

$$\zeta_i = \zeta_i^j \frac{\partial}{\partial q^j}.$$

A long but straight forward calculation yields

$$\begin{aligned} &\left\{ \left(K\left(\frac{\partial}{\partial q^k}, \frac{\partial}{\partial q^j}\right) \dot{q}^j \right)' - M\left(\frac{\partial}{\partial q^k}, \frac{\partial}{\partial q^j}\right) \dot{q}^j - \frac{\partial U}{\partial q^k} \right\} \zeta_i^k \\ &\quad = \{K(\zeta_i, \zeta_j)\theta^j\}' \\ &\qquad - \frac{1}{2}\{K([\zeta_i, \zeta_j], \zeta_k) + K(\zeta_j, [\zeta_i, \zeta_k]) - \zeta_i\{K(\zeta_j, \zeta_k)\}\}\theta^j\theta^k \\ &\qquad - M(\zeta_i, \zeta_j)\theta^j + \zeta_i U = 0 \end{aligned}$$

But we also have

$$B^j_k \zeta^k_i = 0, \forall j.$$

By considering the left null space of the matrix with coefficients given by ζ^k_i, we see that we must have

$$\left\{ K\left(\frac{\partial}{\partial q^k}, \frac{\partial}{\partial q^j}\right) \dot{q}^j \right\}' - M\left(\frac{\partial}{\partial q^k}, \frac{\partial}{\partial q^j}\right) \dot{q}^j - \frac{\partial U}{\partial q^k} = \lambda_j B^j_k.$$

for some set of multipliers $\{\lambda_1, \lambda_2, \ldots\}$. But upon reflection, this is equation 3.1. □

8.2 Introducing the Symmetry

Up to this point, we have not used the symmetry. Nor have used our $\boldsymbol{M} \equiv \mathbf{0}$ assumption. We will do so now.

If we further refine our basis for D to be a basis $\{\kappa_1, \kappa_2, \ldots, \gamma_1, \gamma_2, \ldots, \chi_1, \chi_2, \ldots\}$ aligned with the symmetry, as in definition 6.3.1 on page 47, and correspondingly rename the quasi-velocities μ_j according to

$$\dot{q}^i(\tau) = v^j(\tau)\kappa^i_j + \Omega^j(\tau)\gamma^i_j + \varpi^j(\tau)\chi^i_j,$$

and use the group invariance of the potential U, equation 8.1 splits into

$$\begin{aligned}
(8.3)\qquad \{K(\gamma_i, \gamma_j)\Omega^j\}' &- \frac{1}{2}(\mathcal{L}_{\gamma_i} K)(\kappa_j, \kappa_k)v^j v^k \\
&- \frac{1}{2}(\mathcal{L}_{\gamma_i} K)(\gamma_j, \gamma_k)\Omega^j\Omega^k \\
&- (\mathcal{L}_{\gamma_i} K)(\kappa_j, \gamma_k)v^j\Omega^k \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
(8.4)\qquad \{K(\kappa_i, \kappa_j)v^j\}' &- \frac{1}{2}(\mathcal{L}_{\kappa_i} K)(\kappa_j, \kappa_k)v^j v^k \\
&- \frac{1}{2}(\mathcal{L}_{\kappa_i} K)(\gamma_j, \gamma_k)\Omega^j\Omega^k \\
&- (\mathcal{L}_{\kappa_i} K)(\kappa_j, \gamma_k)v^j\Omega^k \\
&+ \kappa_i U = 0.
\end{aligned}$$

while equation 8.2 splits into

$$
\begin{aligned}
(8.5)\qquad \{K(\gamma_i,\gamma_j)\Omega^j\}' &+ \frac{1}{2}\{K([\gamma_i,\kappa_j],\kappa_k) + K(\kappa_j,[\gamma_i,\kappa_k])\}v^j v^k \\
&+ \frac{1}{2}\{K([\gamma_i,\gamma_j],\gamma_k) + K(\gamma_j,[\gamma_i,\gamma_k])\}\Omega^j\Omega^k \\
&+ \{K([\gamma_i,\kappa_j],\gamma_k) + K(\kappa_j,[\gamma_i,\gamma_k])\}v^j\Omega^k \\
&= 0
\end{aligned}
$$

and

$$
\begin{aligned}
(8.6)\qquad \{K(\kappa_i,\kappa_j)v^j\}' &+ \frac{1}{2}\{K([\kappa_i,\kappa_j],\kappa_k) + K(\kappa_j,[\kappa_i,\kappa_k])\}v^j v^k \\
&+ \frac{1}{2}\{K([\kappa_i,\gamma_j],\gamma_k) + K(\gamma_j,[\kappa_i,\gamma_k])\}\Omega^j\Omega^k \\
&+ \{K([\kappa_i,\kappa_j],\gamma_k) + K(\kappa_j,[\kappa_i,\gamma_k])\}v^j\Omega^k \\
&+ \kappa_i(\\
&- \frac{1}{2}\{\kappa_i\{K(\kappa_j,\kappa_k)\}\}v^j v^k \\
&- \frac{1}{2}\{\kappa_i\{K(\gamma_j,\gamma_k)\}\}\Omega^j\Omega^k = 0.
\end{aligned}
$$

8.3 The Vertical Component

We now assume the flatness conditions, as in definition 7.2.1 on page 52. These conditions once again may be expressed either as

$$
\begin{aligned}
K([\gamma_i,\gamma_j],\gamma_k) + K(\gamma_j,[\gamma_i,\gamma_k]) &= 0 \qquad \text{and} \\
K([\gamma_i,\kappa_j],\kappa_k) + K(\kappa_j,[\gamma_i,\kappa_k]) &= 0.
\end{aligned}
$$

or as

$$
\begin{aligned}
(\mathcal{L}_{\gamma_i}K)(\gamma_j,\gamma_k) &= 0 \qquad \text{and} \\
(\mathcal{L}_{\gamma_i}K)(\kappa_j,\kappa_k) &= 0
\end{aligned}
$$

With this assumption, equation 8.3 becomes

$$\{K(\gamma_i,\gamma_k)\Omega^k\}' = (\mathcal{L}_{\gamma_i}K)(\kappa_j,\gamma_k)v^j\Omega^k.$$

This may be written as the PDE

$$v^j\kappa_j\{K(\gamma_i,\gamma_k)\Omega^k\} + \Omega^m\gamma_m\{K(\gamma_i,\gamma_k)\Omega^k\} = (\mathcal{L}_{\gamma_i}K)(\kappa_j,\gamma_k)v^j\Omega^k.$$

If we set $\Omega = \Omega^i\gamma_i$, then Ω is the vertical component of the trajectory. If, as in section 7.5, we impose the requirement that Ω be group invariant, the previous equation simplifies to

$$v^j\kappa_j\{K(\gamma_i,\gamma_k)\Omega^k\} = (\mathcal{L}_{\gamma_i}K)(\kappa_j,\gamma_k)v^j\Omega^k.$$

We are led to eliminate the v^j functions by considering instead the PDE

$$\kappa_j\{K(\gamma_i,\gamma_k)\Omega^k\} = (\mathcal{L}_{\gamma_i}K)(\kappa_j,\gamma_k)\Omega^k, \tag{8.7}$$

which is the same as

$$\kappa_j\{K(\gamma_i,\Omega)\} = (\mathcal{L}_{\gamma_i}K)(\kappa_j,\Omega). \tag{8.8}$$

We recognize at this point that equation 8.8 is the *adjoint equation*, equation 7.9 in section 7.5. As in section 7.5, this PDE may also be written

$$\kappa_j\{K(\gamma_i,\Omega)\} + K([\gamma_i,\kappa_j],\Omega) + K(\kappa_j,[\gamma_i,\Omega]) = 0, \tag{8.9}$$

which of course is identical to equation 7.10.

In the elementary case (definition 7.3.1 on page 54), equation 8.8 becomes a linear ordinary differential equation, in the same manner that equation 7.6 did (see the discussion leading to proposition 7.3.2 on page 54).

This ODE is precisely the ODE found in the classical analyses of the rolling disk, the ball on a surface of revolution and the rolling axially symmetric body, as we shall show later.

If $q = q(t)$ is the trajectory of the system, we split $\dot{q}$ into its H and S components, as

$$\dot{q} = \dot{q}_H + \dot{q}_S$$

with $\dot{q}_H \in H$ and $\dot{q}_S \in S$. If t_0 is a point in the domain of $q(t)$, then Ω must satisfy the initial conditions

$$\Omega[q(t_0)] = \dot{q}_S(t_0),$$

and we will then have

$$\dot{q}_S(t) = \Omega \circ q(t), \forall t.$$

We obtain one more property of Ω, used in the next subsection, by using equation 8.9 to calculate

$$\begin{aligned}
\kappa_j(K(\Omega,\Omega)) &= (\kappa_j\Omega^i)K(\gamma_i,\Omega) + \Omega^i\kappa_j\{K(\gamma_i,\Omega)\} \\
&= (\kappa_j\Omega^i)K(\gamma_i,\Omega) - \Omega^i\{K([\gamma_i,\kappa_j],\Omega) + K(\kappa_j,[\gamma_i,\Omega])\} \\
&= (\kappa_j\Omega^i)K(\gamma_i,\Omega) - K(\Omega^i[\gamma_i,\kappa_j],\Omega) - K(\kappa_j,\Omega^i[\gamma_i,\Omega])\} \\
&= (\kappa_j\Omega^i)K(\gamma_i,\Omega) - K([\Omega,\kappa_j] + (\kappa_j\Omega^i)\gamma_i,\Omega) \\
&\quad -K(\kappa_j,[\Omega,\Omega] + (\Omega\Omega^i)\gamma_i) \\
&= K([\kappa_j,\Omega],\Omega).
\end{aligned}$$

Applying this to

$$(\mathcal{L}_{\kappa_i}K)(\Omega,\Omega) = \kappa_i\{K(\Omega,\Omega)\} - 2K([\kappa_i,\Omega],\Omega)$$

we have

$$(8.10) \qquad (\mathcal{L}_{\kappa_i}K)(\Omega,\Omega) \;=\; -\kappa_i\{K(\Omega,\Omega)\}.$$

8.4 Reduction to Q/G

In this section, we suppose a group invariant S-valued vector field Ω exists on $Q - \Lambda^{-1}(P)$, such that the trajectory of our nonholonomic system satisfies $\dot{q} = \Omega \circ q$. As we have seen, this supposition holds in the elementary case (proposition 7.3.2 on page 54 and section 7.4).

Now consider equation 8.4. This becomes

$$\begin{aligned}
\{K(\kappa_i,\kappa_j)v^j\}' &- \frac{1}{2}(\mathcal{L}_{\kappa_i}K)(\kappa_j,\kappa_k)v^jv^k \\
&- \frac{1}{2}(\mathcal{L}_{\kappa_i}K)(\Omega,\Omega) \\
&- (\mathcal{L}_{\kappa_i}K)(\kappa_j,\Omega)v^j \\
&+ \kappa_i U \\
= 0.&
\end{aligned}$$

Using equation 8.10, equation 8.4 may be further rearranged to obtain

$$(8.11) \qquad \begin{aligned}
\{K(\kappa_i,\kappa_j)v^j\}' &- \frac{1}{2}(\mathcal{L}_{\kappa_i}K)(\kappa_j,\kappa_k)v^jv^k \\
&- (\mathcal{L}_{\kappa_i}K)(\kappa_j,\Omega)v^j \\
&+ \kappa_i\left\{U + \frac{1}{2}K(\Omega,\Omega)\right\} = 0.
\end{aligned}$$

Lemma 8.4.1. *A function $\overline{U} : Q/G - P \to \Re$ is determined by the requirement*

$$\overline{U} \circ \Lambda = U + \frac{1}{2}K(\Omega, \Omega).$$

Proof. The function $U + \frac{1}{2}K(\Omega, \Omega) : Q - \Lambda^{-1}(P) \to \Re$ is group invariant. □

Definition 8.4.1. *The function $\overline{U}$ on Q/G, in the preceding lemma, is called the* reduced potential.

Lemma 8.4.2. *A metric $\overline{K}$ on Q/G is induced by the requirement*

$$\overline{K}(T\Lambda(\tau), T\Lambda(\tau)) = K(\tau, \tau), \forall \tau \in H \dotplus R.$$

Proof. Recall that $T\Lambda(H \dotplus R) = T(Q/G)$ holds, and that $T\Lambda(\tau) = 0$ with $\tau \in H \dotplus R$ implies $\tau = 0$ (refer to the comments following proposition 6.2.1 on page 46). And K is group invariant. □

Definition 8.4.2. *The metric $\overline{K}$, in the preceding lemma, is called the* reduced kinetic inner product.

Lemma 8.4.3. *For each i, a vector field $\overline{\kappa}_i$ on (the appropriate portion of) Q/G is induced by the requirement*

$$\overline{\kappa}_i \circ \Lambda = T\Lambda \circ \kappa_i.$$

Proof. Each vector field κ_i is group invariant. □

If we put $\overline{D} = T\Lambda(D)$, then $\overline{D}$ is a distribution on Q/G, $\{\overline{\kappa}_1, \overline{\kappa}_2, \ldots\}$ is a local basis for $\overline{D}$, and we may obtain any local basis for $\overline{D}$ in this way.

Lemma 8.4.4. *The following equality holds:*

$$(\mathcal{L}_{\kappa_i} K)(\kappa_j, \kappa_k) = \{(\mathcal{L}_{\overline{\kappa}_i} \overline{K})(\overline{\kappa}_j, \overline{\kappa}_k)\} \circ \Lambda$$

Proof. We calculate

$$\begin{aligned}
(\mathcal{L}_{\kappa_i} K)(\kappa_j, \kappa_k) &= \kappa_i\{K(\kappa_j, \kappa_k)\} - K([\kappa_i, \kappa_j], \kappa_k) - K(\kappa_j, [\kappa_i, \kappa_k]) \\
&= \kappa_i\{\overline{K}(\overline{\kappa}_j \circ \Lambda, \overline{\kappa}_k \circ \Lambda)\} \\
&\quad -\overline{K}(T\Lambda([\kappa_i, \kappa_j]), \overline{\kappa}_k \circ \Lambda) \\
&\quad -\overline{K}(\overline{\kappa}_j \circ \Lambda, T\Lambda([\kappa_i, \kappa_k])) \\
&= (\overline{\kappa}_i \circ \Lambda)\{\overline{K}(\overline{\kappa}_j, \overline{\kappa}_k)\} \\
&\quad -\overline{K}([\overline{\kappa}_i \circ \Lambda, \overline{\kappa}_j \circ \Lambda], \overline{\kappa}_k \circ \Lambda) \\
&\quad -\overline{K}(\overline{\kappa}_j \circ \Lambda, [\overline{\kappa}_i \circ \Lambda, \overline{\kappa}_k \circ \Lambda]) \\
&= \{(\mathcal{L}_{\overline{\kappa}_i} \overline{K})(\overline{\kappa}_j, \overline{\kappa}_k)\} \circ \Lambda.
\end{aligned}$$

□

Lemma 8.4.5. *There is a group invariant 2-form M on Q such that, for any group invariant vector fields τ and ρ with $\tau, \rho \in H$,*

$$M(\tau, \rho) = (\mathfrak{L}_\tau K)(\rho, \Omega)$$

holds. A 2-form $\overline{M}$ on Q/G is induced by the requirement

$$\overline{M}(T\lambda(\tau), T\lambda(\rho)) = M(\tau, \rho), \ \forall \tau \in H \dotplus R.$$

Proof. Recall that $T\lambda(H \dotplus R) = T(Q/G)$ holds. Therefore we need only to define M.

Let τ and ρ be group invariant vector fields with $\tau, \rho \in H$. We have

$$\begin{aligned}
(\mathfrak{L}_\tau K)(\rho, \Omega) &= \tau\{K(\rho, \Omega)\} - K([\tau, \rho], \Omega) - K(\rho, [\tau, \Omega]) \\
&= -K([\tau, \rho], \Omega) - K(\rho, [\tau, \Omega]).
\end{aligned}$$

But each of these terms is linear, and anti-symmetric in τ and ρ. To see this, we have locally

$$\begin{aligned}
K([\tau, \rho], \Omega) &= K([\tau^i \kappa_i, \rho^j \kappa_j], \Omega) \\
&= \tau^i \rho^j \{K([\kappa_i, \kappa_j], \Omega)\} + (\tau \rho^j) K(\kappa_j, \Omega) - (\rho \tau^i) K(\kappa_i, \Omega) \\
&= \tau^i \rho^j \{K([\kappa_i, \kappa_j], \Omega)\}
\end{aligned}$$

and

$$\begin{aligned}
K(\rho, [\tau, \Omega]) &= \rho^i K(\kappa_i, [\tau^j \kappa_j, \Omega^k \gamma_k]) \\
&= \rho^i \tau^j \{\Omega^k K(\kappa_i, [\kappa_j, \gamma_k])\} \\
&\quad + \rho^i K(\kappa_i, (\tau \Omega^k) \gamma_k) - \rho^i K(\kappa_i, (\Omega \tau^j) \kappa_j) \\
&= \rho^i \tau^j \{\Omega^k K(\kappa_i, [\kappa_j, \gamma_k])\} - \rho^i (\Omega \tau^j) K(\kappa_i, \kappa_j) \\
&= \rho^i \tau^j \{\Omega^k K(\kappa_i, [\kappa_j, \gamma_k])\}.
\end{aligned}$$

These are bilinear expressions. The coefficent in the first is clearly anti-symmetric. The coefficient in the second is anti-symmetric due to one of the flatness conditions, equation 7.3 on page 51.

This serves to define the value M should take when restricted to vectors in H. The group invariance of the coefficients, in the local expressions, show that this definition is group invariant.

But the inner product K may be used to project any vector to H. Using this we extend the definition of M all of TQ. The group invariance of K ensures the group invariance of M. □

In light of the preceding lemmas, it is now apparent that equation 8.11 pushes down to

$$\begin{aligned} \{\overline{K}(\overline{\kappa}_i, \overline{\kappa}_j)v^j\}' &- \frac{1}{2}(\mathcal{L}_{\overline{\kappa}_i}\overline{K})(\overline{\kappa}_j, \overline{\kappa}_k)v^j v^k \\ &- \overline{M}(\overline{\kappa}_i, \overline{\kappa}_j)v^j + \overline{\kappa}_i \overline{U} = 0 \end{aligned}$$

on Q/G. These reduced equations on Q/G are in the same form as equation 8.1. Observe that the *reduced magnetic field form* $\overline{M}$ need not be closed.

Using lemma 8.1.1, we have:

Proposition 8.4.1. *Assuming*

- *the flatness conditions are satisfied and*
- *the adjoint equation is integrable,*

the projection onto Q/G of the trajectory on Q, corresponds to the trajectory of a reduced nonholonomic system on Q/G, with constraint distribution $\overline{D}$, magnetic field form $\overline{M}$ and natural Lagrangian $\overline{L} : T(Q/G) \to \mathbb{R}$ given by

$$\overline{L}(\tau) = \frac{1}{2}\overline{K}(\tau, \tau) - \overline{U} \circ \pi_{Q/G}(\tau),$$

where $\pi_{Q/G} : T(Q/G) \to Q/G$ is the tangent bundle projection.

Note that $K(\dot{q}) + U \circ \pi(\dot{q}) = \overline{K} \circ T\Lambda(\dot{q}) + \overline{U} \circ \pi_{Q/G} \circ T\Lambda(\dot{q})$ holds, so that the energy in the reduced system is the same as that for the original.

Corollary 8.4.1. *If in addition to the requirements of proposition 8.4.1, the reduced magnetic field form is closed, and the dimension assumption, $\overline{D} = T(Q/G)$, is satisfied, the reduced motion on Q/G is Hamiltonian.*

Finally, note that if the dimension of Q/G is 1, then $\overline{M}$ must be zero. If the dimension of Q/G is 2, then $\overline{M}$ is necessarily closed.

Example. Nonholonomically Constrained Particle in $\mathbb{R}^3$

Recall that the kinetic inner product for this example is given by

$$K((\dot{x}, \dot{y}, \dot{z}), (\underline{\dot{x}}, \underline{\dot{y}}, \underline{\dot{z}})) = \dot{x}\underline{\dot{x}} + \dot{y}\underline{\dot{y}} + \dot{z}\underline{\dot{z}},$$

with a basis aligned with the symmetry given by

$$\begin{aligned} \kappa_1 &= (0, 1, 0) \qquad \text{and} \\ \gamma_1 &= (1, 0, y). \end{aligned}$$

We also have previously found

$$\Omega = \frac{C}{\sqrt{1+y^2}}(1, 0, y),$$

where C is a constant of integration. The reduced potential becomes

$$\begin{aligned} \overline{U} &= \frac{1}{2}K(\Omega, \Omega) \\ &= \frac{1}{2}\frac{C^2}{1+y^2}(1+y^2) \\ &= \frac{1}{2}C^2, \end{aligned}$$

while the reduced kinetic energy is

$$\begin{aligned} \overline{K} &= \frac{1}{2}K(\kappa_1, \kappa_1)\dot{y}^2 \\ &= \frac{1}{2}\dot{y}^2. \end{aligned}$$

In this case Q/G is parameterized by y, and has dimension 1. So there is no magnetic term. The reduced equation of motion will be

$$\ddot{y} = 0.$$

□

8.5 A Helpful Proposition

In the following sections, we will examine a number of examples. In each case, we will want to construct a group invariant basis for D. All of the cases considered will satisfy the criteria of the next proposition.

Proposition 8.5.1. *Let the configuration manifold be a cross product of two other manifolds, $Q = M \times P$. Let $\Sigma : Q \to P$ be the associated projection map, and suppose that*

- $T\Sigma(D) = TP$ *and*
- $T\Sigma(\tau) = 0$ *implies* $\tau = 0$

hold. Then given a vector field ξ on P, there is a unique vector field ξ^D on Q with $\xi^D \in D$ and $T\Sigma \circ \xi^D = \xi \circ \Sigma$ (called the lift *of ξ). Suppose also that the symmetry group G for Q has an action on P alone such that*

- $\Sigma \circ l_g = l_g^P \circ \Sigma, \ \forall g \in G$

holds (l^P refers to the action on P). If ξ is group invariant under the group action on P, then ξ^D is group invariant under the group action on Q.

Proof. The existence of a unique solution of $T\Sigma \circ \xi^D = \xi \circ \Sigma$ at each point of Q is clear. The proof that this solution is differentiable is tedious, and typical of proofs in differential geometry. We will not provide this here.

If in addition ξ is group invariant ($Tl_g^P \circ \xi = \xi \circ l_g^P, \ \forall g \in G$), then we have

$$\begin{aligned}
T\Sigma \circ Tl_g \circ \xi^D &= T(\Sigma \circ l_g) \circ \xi^D \\
&= T(l_g^P \circ \Sigma) \circ \xi^D \\
&= Tl_g^P \circ T\Sigma \circ \xi^D \\
&= Tl_g^P \circ \xi \circ \Sigma \\
&= \xi \circ l_g^P \circ \Sigma \\
&= \xi \circ \Sigma \circ l_g \\
&= T\Sigma \circ \xi^D \circ l_g.
\end{aligned}$$

From this we have $Tl_g \circ \xi^D = \xi^D \circ l_g$. So ξ^D is group invariant. $\square$

Chapter 9

Axially Symmetric Rolling Body

In this section we consider the example of an axially symmetric body rolling without slipping on a horizontal plane in the presence of a uniform gravitational field.

9.1 Formulating the Problem

In this subsection we formulate the Lagrangian and the constraint equations for this system, using globally defined quasi-velocities.

Consider figure 9.1 below.

The figure is a cross-section. The point C is the center of mass. The point A is the point of contact of the body with the horizontal plane. The line BC is the axis of symmetry of the body. The line AB is perpendicular to the horizontal plane.

The angle θ is as shown. We regard each of the distances a, r, h and d as an even function of θ, with period 2π. These distances are related by

$$\begin{aligned} h(\theta) &= r(\theta) + a(\theta)\cos\theta \qquad \text{and} \\ d(\theta) &= a(\theta) + r(\theta)\cos\theta. \end{aligned}$$

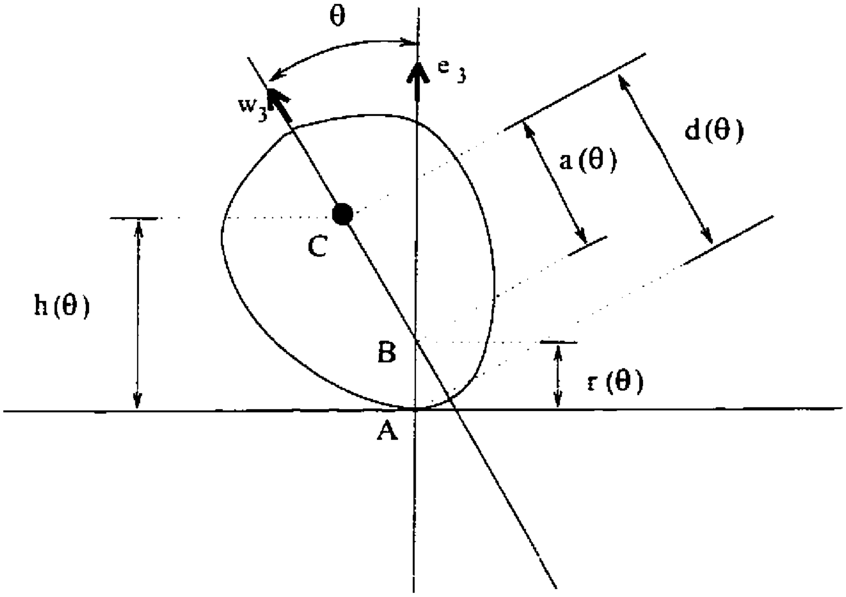


Figure 9.1: Axially Symmetric Rolling Body

which may be solved to obtain

$$r(\theta) = \frac{h(\theta) - d(\theta)\cos\theta}{\sin^2\theta} \quad \text{and}$$
$$a(\theta) = \frac{d(\theta) - h(\theta)\cos\theta}{\sin^2\theta}.$$

We assume that h and d are defined everywhere, are continuous, and have continuous first derivatives.

By considering the appearance of the figure as θ varies, we see that we also need to assume that h and d satisfy

$$h(\theta) \geq 0, \forall\theta,$$
$$h(0) = d(0) \quad \text{and}$$
$$h(\pi) = -d(\pi).$$

The behaviour of a and r at multiples of π may vary.

Example. The Rolling Disk

For a rolling disk, multiples of π are not valid values for θ, and we have $a(\theta) = -\rho\cot\theta$, $r(\theta) = \rho\csc\theta$, $h(\theta) = \rho\sin\theta$ and $d(\theta) = 0$, $\forall\theta \in (0,\pi)$, where ρ is the radius of the disk. □

Example. Routh's Model for a Top

For a *sphere*, a and r are constant. For an axially symmetric body with a spherical base, we may have a greater than r. This model was used by Jellett [22], Routh [31] and Gallop [16] as a model for a *top*. Allowing $r = 0$, it is apparent intuitively that the base must be a point, and we should obtain the *Lagrange top*. This will be seen to be so below. □

We will use the ideas and notation of section 2.5 in this section.

We take the fixed axes so that e_3 points upwards in the figure. We take the moving axes and their associated orthonormal vectors, w_1, w_2 and w_3, so that w_3 points along the axis of symmetry of the body, in the direction from B to C. The position of the center of mass be given with respect to the fixed axes by

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 e_1 + x_2 e_2 + x_3 e_3.$$

The orientation of the body is given by $W \in SO(3)$ with

$$W = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}.$$

The angular velocity is given by ν with respect to the fixed axes, and by ω with respect to the body (moving) axes. The total mass of the body is given by m. The inertia matrix of the body with respect to the body axes is given by J. The choice of body axes ensures that J is diagonal.

$$J = \begin{bmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_3 \end{bmatrix}.$$

and that $J_1 = J_2$ holds. The kinetic energy of the body is given by

$$K = \frac{1}{2}\omega^T J \omega + \frac{1}{2} m \dot{x}^T \dot{x}.$$

The potential energy of the body is given by

$$U = m g e_3^T x.$$

where g is the gravitational constant.

The Lagrangian for this system is therefore given by

$$L = \frac{1}{2}\omega^T J\omega + \frac{1}{2}m\dot{x}^T\dot{x} - mge_3^T x. \tag{9.1}$$

The *rolling condition* is that the velocity of the point on the body that is in contact with the horizontal plane has zero velocity. Consider once again figure 9.1. The vector from the point C to A is $-(aw_3+r\epsilon_3)$. Using equation 2.8 on page 17, the constraint then is

$$\begin{aligned} 0 &= \dot{x} - A(\nu)(aw_3 + r\epsilon_3) \\ &= \dot{x} - A(W\omega)(aW\epsilon_3 + r\epsilon_3) \\ &= \dot{x} + A(aW\epsilon_3 + r\epsilon_3)W\omega \end{aligned} \tag{9.2}$$

The angle θ in figure 9.1 will be a function of W, with values in the interval $[0, \pi]$. In fact we have

$$\cos\theta = \epsilon_3^T w_3 = \epsilon_3^T W\epsilon_3$$

and

$$\sin^2\theta = \{A(\epsilon_3)W\epsilon_3\}^T A(\epsilon_3)W\epsilon_3.$$

Note that using x and W as coordinates suggests that the configuration space is $SO(3) \times \Re^3$. However, this is not really so, due to the constraint

$$h(\theta(W)) = \epsilon_3^T x. \tag{9.3}$$

The actual configuration manifold Q is therefore a submanifold in $SO(3)\times\Re^3$, easily seen to be diffeomorphic to $SO(3) \times \Re^2$.

Also, equation 9.2 suggests that there are three linear velocity constraints. But one of these simply requires that the tangent vector to a trajectoriy lies in TQ. There are in fact only two linear nonholonomic constraints in TQ.

So the configuration manifold Q is 5-dimensional, and the constraint distribution D within TQ is 3-dimensional.

Example. The Lagrange Top

Finally, observe that if $r = 0$ holds with a constant, then equation 9.2 is just

$$\begin{aligned} 0 &= \dot{x} + A(aW\epsilon_3)W\omega \\ &= \dot{x} + aWA(\epsilon_3)\omega \\ &= \dot{x} - aWA(\omega)\epsilon_3 \\ &= \{x - aW\epsilon_3\}'. \end{aligned}$$

So, choosing the origin for the fixed axes to be at the point of contact, we obtain $x = aWe_3$. This was apparent from figure 9.1. This is simply the case of the *Lagrange top*. □

9.2 The Group Symmetry

In this section we describe a group symmetry of this system associated with the Lie group

$$G = SE(2) \times SO(2).$$

Recall now that the Lie group $SE(3)$ is the manifold $SO(3) \times \Re^3$ with the group product given by

$$(H, y)(K, z) = (HK, y + Hz).$$

It will be convenient for us to think of $SO(2)$ as the Lie subgroup $\{R_3(\phi) | \phi \in \Re\}$ of $SO(3)$, and to think of $SE(2)$ as the Lie subgroup $\{(R_3(\phi), (\xi, \zeta, 0)) | \phi, \xi, \zeta \in \Re\}$ of $SE(3)$.

To describe the group action, we parameterize G (that is, $SE(2) \times SO(2)$) by $\Re^4$ using the map

$$(\phi, \xi, \zeta, \psi) \mapsto ((R_3(\phi), (\xi, \zeta, 0)), R_3(\psi)).$$

The associated group action on $SO(3) \times \Re^3$ is given by the map $(W, x) \mapsto (\overline{W}, \overline{x})$, with

$$\overline{W} = R_3(\phi) W R_3(\psi)^T \tag{9.4}$$
$$\overline{x} = R_3(\phi) x + \xi e_1 + \zeta e_2. \tag{9.5}$$

We observe first that

$$e_3^T \overline{x} = e_3^T x$$

holds. Thinking of θ as a function on $SO(3) \times \Re^3$, we also have

$$\begin{aligned} \cos \overline{\theta} &= e_3^T \overline{W} e_3 = e_3^T R_3(\phi) W R_3(\psi)^T e_3 \\ &= e_3^T W e_3 = \cos \theta, \end{aligned}$$

and so obtain

$$\overline{\theta} = \theta.$$

Lemma 9.2.1. *The group action of G on $SO(3) \times \mathbb{R}^3$, induces an action on the configuration manifold Q.*

Proof. Q is determined by the constraint $h(\theta) = e_3^T x$, which is preserved by the group action, since $\overline{\theta} = \theta$ and $e_3^T \overline{x} = e_3^T x$ hold. □

The following lemma states a well known fact, which we will use below.

Lemma 9.2.2. *The map $\mathbb{R}^3 \to SO(3) : (\phi, \theta, \psi) \mapsto R_3(\phi) R_1(\theta) R_3(\psi)$ is surjective. The parameters ϕ, θ and ψ are usually referred to as* Euler angles.

Proposition 9.2.1. *The function induced on Q/G by $\theta : Q \to [0, \pi]$ is bijective. Hence Q/G is a line segment.*

Proof. Suppose (W_1, z_1) and (W_2, z_2) are in Q, and that we have

$$\theta(W_1) = \theta(W_2).$$

Then we must have

$$e_3^T W_1 e_3 = e_3^T W_2 e_3.$$

Let

$$W_1 = R_3(\phi_1) R_1(\theta_1) R_3(\psi_1)$$

and

$$W_2 = R_3(\phi_2) R_1(\theta_2) R_3(\psi_2)$$

hold. We must then have $e_3^T R_1(\theta_1) e_3 = e_3^T R_1(\theta_2) e_3$, which implies $\cos\theta_1 = \cos\theta_2$. So we have either

$$\begin{aligned} R_1(\theta_2) &= R_1(\theta_1) \qquad\qquad \text{or} \\ R_1(\theta_2) &= R_1(-\theta_1). \end{aligned}$$

The first case yields $W_2 = R_3(\phi_2 - \phi_1) W_1 R_3(\psi_1 - \psi_2)^T$. Otherwise we may use the identity

$$R_1(-\varphi) = R_3(\pi) R_1(\varphi) R_3(\pi)$$

to yield $W_2 = R_3(\phi_2 - \phi_1 + \pi) W_1 R_3(\psi_1 - \psi_2 - \pi)^T$. We may write either of these results as

$$W_2 = R_3(\phi) W_1 R_3(\psi)^T.$$

Next, $\epsilon_3^T z_1 = \epsilon_3^T z_2$ must hold on Q, and so we also have

$$\epsilon_3^T \{z_2 - R_3(\phi) z_1\} = 0.$$

We conclude that (W_1, z_1) and (W_2, z_2) are in the same group orbit. $\square$

We must show that the Lagrangian L is preserved by this group action. Once again we have

$$L = \frac{1}{2}\omega^T J \omega + \frac{1}{2} m \dot{x}^T \dot{x} - mg\epsilon_3^T x.$$

We immediately have

$$\epsilon_3^T \bar{x} = \epsilon_3^T x$$

and

$$\dot{\bar{x}} = R_3(\phi)\dot{x}.$$

So the last two terms are preserved.

From equation 9.4 we have

$$\begin{aligned}
\bar{W} A(\bar{\omega}) &= \dot{\bar{W}} \\
&= R_3(\phi) \dot{W} R_3(\psi)^T \\
&= R_3(\phi) W A(\omega) R_3(\psi)^T \\
&= R_3(\phi) W R_3(\psi)^T A(R_3(\psi)\omega) \\
&= \bar{W} A(R_3(\psi)\omega),
\end{aligned}$$

so that

$$\bar{\omega} = R_3(\psi)\omega$$

holds, which leads to

$$\begin{aligned}
\frac{1}{2}\bar{\omega}^T J \bar{\omega} &= \frac{1}{2}\omega R_3(\psi)^T J R_3(\psi)\omega \\
&= \frac{1}{2}\omega^T J \omega.
\end{aligned}$$

since $J_1 = J_2$ holds.

Thus we obtain $\overline{L} = L$, and so L is preserved by the group action.

We must also show that the constraint.

$$0 = \dot{x} + A(aW\epsilon_3 + r\epsilon_3)W\omega.$$

is preserved. For this we have

$$\begin{aligned} A(a\overline{W}\epsilon_3 + r\epsilon_3)\overline{W}\overline{\omega} &= A(aR_3(\phi)WR_3(\psi)^T\epsilon_3 + r\epsilon_3)R_3(\phi)W\omega \\ &= R_3(\phi)A(aW\epsilon_3 + r\epsilon_3)W\omega \end{aligned}$$

and then

$$\dot{\overline{x}} + A(a\overline{W}\epsilon_3 + r\epsilon_3)\overline{W}\overline{\omega} = R_3(\phi)\{\dot{x} + A(aW\epsilon_3 + r\epsilon_3)W\omega\}.$$

Since $R_3(\phi)$ is non-singular, we see that the constraint is preserved by the group action.

9.3 A Basis Aligned with the Symmetry

We now apply the theory of chapter 6 to the current group action, given by equation 9.4 and equation 9.5, which we repeat as

$$\begin{aligned} \overline{W} &= R_3(\phi)WR_3(\psi)^T \\ \overline{x} &= R_3(\phi)x + \xi\epsilon_1 + \zeta\epsilon_2. \end{aligned}$$

To find a basis of vector fields for the distribution V, of vectors tangent to group orbits, we take partial derivatives. Each vector field obtained in this way will be the fundamental vector field associated with an element of the Lie algebra of $SE(2) \times SO(2)$, due to the way the parameterization was chosen.

Taking partial derivatives, we obtain

$$
\begin{aligned}
\left.\frac{\partial(\overline{W,x})}{\partial\xi}\right|_{(0,0,0,0)} &= (0, e_1), \\
\left.\frac{\partial(\overline{W,x})}{\partial\zeta}\right|_{(0,0,0,0)} &= (0, e_2), \\
\left.\frac{\partial(\overline{W,x})}{\partial o}\right|_{(0,0,0,0)} &= (A(e_3)W, A(e_3)x) \\
&= (WA(W^T e_3), A(e_3)x) \qquad \text{and} \\
\left.\frac{\partial(\overline{W,x})}{\partial\iota}\right|_{(0,0,0,0)} &= (-WA(e_3), 0) \\
&= (WA(-e_3), 0).
\end{aligned}
$$

We may map these to $(\omega, \dot{x})$-space to obtain instead

$$
\begin{aligned}
\frac{\partial}{\partial\xi} &\mapsto (0, e_1), \\
\frac{\partial}{\partial\zeta} &\mapsto (0, e_2), \\
\frac{\partial}{\partial o} &\mapsto (W^T e_3, A(e_3)x) \qquad \text{and} \\
\frac{\partial}{\partial\iota} &\mapsto (-e_3, 0).
\end{aligned}
$$

Each of these vector fields will necessarily take values in TQ when evaluated on Q. These four vectors are independent except when $W^T e_3 = \pm e_3$ holds. But this condition implies $e_3^T W e_3 = \pm 1$, so that W projects to an end point of Q/G. We would expect the basis to collapse at such points.

To find a basis of vector fields for the constraint distribution D, in $(\omega, \dot{x})$-space, we simply substitute values for ω in the constraint,

$$0 = \dot{x} + A(aWe_3 + re_3)W\omega.$$

Our choices for ω are made with proposition 8.5.1 on page 69 in mind.

For $\omega = e_3$ we obtain

$$
\begin{aligned}
\dot{x} &= -A(aWe_3 + re_3)We_3 \\
&= -rA(e_3)We_3.
\end{aligned}
$$

Using this we have

$$\begin{aligned} K(\kappa, \gamma_1) &= [A(W^T e_3)e_3]^T J e_3 \\ &= J_3[A(W^T e_3)e_3]^T e_3 \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} K(\kappa, \gamma_2) &= [A(W^T e_3)e_3]^T J W^T e_3 \\ &= J_1[A(W^T e_3)e_3]^T W^T e_3 \\ &= 0. \end{aligned}$$

We have used

$$J e_3 = J_3 e_3$$

for the first calculation above, and

$$J\{A(W^T e_3)e_3\} = J_1\{A(W^T e_3)e_3\}$$

for the second. Hence $\{\kappa\}$ is a basis for $H = D \cap S^\perp$.

We have not as yet shown that γ_1, γ_2 and κ are group invariant vector fields. We defer this to the next subsection, in which we also show that this basis satisfies the flatness conditions.

9.4 Group Invariance and the Flatness Conditions

The calculations in this subsection are based on the following notion. Thinking of γ_1, γ_2 and κ as differentiation operators, we have

(9.6)
$$\begin{aligned} \gamma_1 W &= WA(e_3), \\ \gamma_1 x &= -rA(e_3)We_3, \\ \gamma_2 W &= WA(W^T e_3) \\ &= A(e_3)W, \\ \gamma_2 x &= aA(e_3)We_3, \\ \kappa W &= WA(A(W^T e_3)e_3) \\ &= A(A(e_3)We_3)W \qquad \text{and} \\ \kappa x &= -de_3 + hWe_3. \end{aligned}$$

To apply these differentiation operators to general expressions involving W and x, we must apply the usual rules for taking derivatives of products and sums. In particular, we may find

$$\begin{aligned}
\gamma_1(\cos\theta) &= \gamma_1(\epsilon_3^T W \epsilon_3) \\
&= \epsilon_3^T(\gamma_1 W)\epsilon_3 \\
&= \epsilon_3^T W A(\epsilon_3)\epsilon_3 \\
&= 0 \qquad\qquad \text{and} \\
\gamma_2(\cos\theta) &= \epsilon_3^T A(\epsilon_3) W \epsilon_3 \\
&= 0,
\end{aligned}$$

from which we conclude that

$$\gamma_1\theta = \gamma_2\theta = 0$$

holds. We may also calculate

$$\begin{aligned}
-\sin\theta(\kappa\theta) &= \kappa(\cos\theta) \\
&= \kappa(\epsilon_3^T W \epsilon_3) \\
&= \epsilon_3^T A(A(\epsilon_3) W \epsilon_3) W \epsilon_3 \\
&= -\epsilon_3^T A(W\epsilon_3) A(\epsilon_3) W \epsilon_3 \\
&= -\{A(\epsilon_3) W \epsilon_3\}^T A(\epsilon_3) W \epsilon_3 \\
&= -\sin^2\theta,
\end{aligned}$$

so that

$$\kappa\theta = \sin\theta$$

holds.

Proposition 9.4.1. *The vector fields γ_1, γ_2 and κ are group invariant.*

Proof. For γ_1 we calculate

$$\begin{aligned}
\gamma_1\overline{W} &= \gamma_1(R_3(\phi)WR_3(\psi)^T) \\
&= R_3(\phi)WA(e_3)R_3(\psi)^T \\
&= R_3(\phi)WR_3(\psi)^TA(R_3(\psi)e_3) \\
&= \overline{W}A(e_3) \qquad\qquad \text{and} \\
\gamma_1\overline{x} &= \gamma_1(R_3(\phi)x + \xi e_1 + \zeta e_2) \\
&= R_3(\phi)(-rA(e_3)We_3) \\
&= -rA(R_3(\phi)e_3)R_3(\phi)WR_3(\psi)^Te_3 \\
&= -rA(e_3)\overline{W}e_3.
\end{aligned}$$

For γ_2 we calculate

$$\begin{aligned}
\gamma_2\overline{W} &= \gamma_2(R_3(\phi)WR_3(\psi)^T) \\
&= R_3(\phi)A(e_3)WR_3(\psi)^T \\
&= A(R_3(\phi)e_3)R_3(\phi)WR_3(\psi) \\
&= A(e_3)\overline{W} \qquad\qquad \text{and} \\
\gamma_2\overline{x} &= \gamma_2(R_3(\phi)x + \xi e_1 + \zeta e_2) \\
&= R_3(\phi)(aA(e_3)We_3) \\
&= aA(e_3)\overline{W}e_3.
\end{aligned}$$

For κ we calculate

$$\begin{aligned}
\kappa\overline{W} &= \kappa(R_3(\phi)WR_3(\psi)^T) \\
&= R_3(\phi)A(A(e_3)We_3)WR_3(\psi)^T \\
&= A(A(R_3(\phi)e_3)R_3(\phi)We_3)R_3(\phi)WR_3(\psi)^T \\
&= A(A(e_3)R_3(\phi)WR_3(\psi)^Te_3)R_3(\phi)WR_3(\psi)^T \\
&= A(A(e_3)\overline{W}e_3)\overline{W} \qquad\qquad \text{and} \\
\kappa\overline{x} &= \kappa(R_3(\phi)x + \xi e_1 + \zeta e_2) \\
&= R_3(\phi)(-de_3 + hWe_3) \\
&= -de_3 + hR_3(\phi)WR_3(\psi)^Te_3 \\
&= -de_3 + h\overline{W}e_3.
\end{aligned}$$

These results may now be compared with equations 9.6.

In light of proposition 8.5.1, it would have been sufficient in this proof to have only calculated the effects of the vectors on $\overline{W}$, but little effort would have been saved. $\square$

Next we find the commutators of these vector fields.

Lemma 9.4.1. *The pair-wise brackets of γ_1, γ_2 and κ (in $(\omega, \dot{x})$-space) are*

$$
\begin{aligned}
(9.7) \qquad [\gamma_1, \gamma_2] &= (0, rA(e_3)A(e_3)W e_3), \\
(9.8) \qquad [\kappa, \gamma_1] &= (0, -(r' \sin\theta + r\cos\theta)A(e_3)W e_3) \qquad \textit{and} \\
(9.9) \qquad [\kappa, \gamma_2] &= (0, (a' \sin\theta - r)A(e_3)W e_3),
\end{aligned}
$$

where we use $r' = dr/d\theta$ and $a' = da/d\theta$.

Proof. To obtain equation 9.7 we calculate

$$
\begin{aligned}
\gamma_1(\gamma_2 x) &= \gamma_1(aA(e_3)W e_3) \\
&= aA(e_3)WA(e_3)e_3 = 0, \\
\gamma_2(\gamma_1 x) &= \gamma_2(-rA(e_3)W e_3) \\
&= -rA(e_3)A(e_3)W e_3, \\
\gamma_1(\gamma_2 W) &= \gamma_1(A(e_3)W) \\
&= A(e_3)WA(e_3) \qquad \text{and} \\
\gamma_2(\gamma_1 W) &= \gamma_2(WA(e_3)) \\
&= A(e_3)WA(e_3).
\end{aligned}
$$

To obtain equation 9.8 we calculate

$$\begin{aligned}
\kappa(\gamma_1 x) &= \kappa(-rA(\epsilon_3)W\epsilon_3) \\
&= -(r'\sin\theta)A(\epsilon_3)W\epsilon_3 - rA(\epsilon_3)A(A(\epsilon_3)W\epsilon_3)W\epsilon_3 \\
&= -(r'\sin\theta)A(\epsilon_3)W\epsilon_3 + rA(\epsilon_3)A(W\epsilon_3)A(\epsilon_3)W\epsilon_3 \\
&= (-r'\sin\theta I + r(W\epsilon_3\epsilon_3^T - (\epsilon_3^T W\epsilon_3)I))A(\epsilon_3)W\epsilon_3 \\
&= -(r'\sin\theta + r\cos\theta)A(\epsilon_3)W\epsilon_3. \\
\gamma_1(\kappa x) &= \gamma_1(-d\epsilon_3 + hW\epsilon_3) \\
&= hWA(\epsilon_3)\epsilon_3 = 0. \\
\kappa(\gamma_1 W) &= \kappa(WA(\epsilon_3)) \\
&= A(A(\epsilon_3)W\epsilon_3)WA(\epsilon_3) \qquad \text{and} \\
\gamma_1(\kappa W) &= \gamma_1(A(A(\epsilon_3)W\epsilon_3)W) \\
&= A(A(\epsilon_3)WA(\epsilon_3)\epsilon_3)W + A(A(\epsilon_3)W\epsilon_3)WA(\epsilon_3) \\
&= A(A(\epsilon_3)W\epsilon_3)WA(\epsilon_3).
\end{aligned}$$

To obtain equation 9.9 we calculate

$$\begin{aligned}
\kappa(\gamma_2 x) &= \kappa(aA(\epsilon_3)W\epsilon_3) \\
&= (a'\sin\theta + a\cos\theta)A(\epsilon_3)W\epsilon_3. \\
\gamma_2(\kappa x) &= \gamma_2(-d\epsilon_3 + hW\epsilon_3) \\
&= hA(\epsilon_3)W\epsilon_3 \\
&= (r + a\cos\theta)A(\epsilon_3)W\epsilon_3. \\
\kappa(\gamma_2 W) &= \kappa(A(\epsilon_3)W) \\
&= A(\epsilon_3)A(A(\epsilon_3)W\epsilon_3)W \qquad \text{and} \\
\gamma_2(\kappa W) &= \gamma_2(A(A(\epsilon_3)W\epsilon_3)W) \\
&= A(A(\epsilon_3)A(\epsilon_3)W\epsilon_3)W + A(A(\epsilon_3)W\epsilon_3)A(\epsilon_3)W \\
&= \{A(\epsilon_3)A(A(\epsilon_3)W\epsilon_3) - A(A(\epsilon_3)W\epsilon_3)A(\epsilon_3)\}W \\
&\quad + A(A(\epsilon_3)W\epsilon_3)A(\epsilon_3)W \\
&= A(\epsilon_3)A(A(\epsilon_3)W\epsilon_3)W.
\end{aligned}$$

$\square$

We may interpret these results geometrically. The bracket of two vector fields is a measure of the commutativity of the flows of these vector fields. A flow which satisfies the constraints is normally called a *virtual motion* in mechanics.

For the virtual motion associated with γ_1, the body is rotated about its axis of symmetry. The body rolls while maintaining the orientation of its symmetry axis. In figure 9.1, the body is moving perpendicular to the page. The cross-section in the figure remains parallel to the page, and looks the same, while moving with the body.

For γ_2, the body is rotated about the vertical axis through the point of contact with the plane. The point of contact remains fixed.

For κ, the body is rolled while keeping the axis of symmetry of the body, and the vertical axis through the point of contact, in a constant plane. In figure 9.1, θ changes, and the point of contact with the plane will move to one side.

Lemma 9.4.1 asserts that if two of these virtual motions are applied to the body, the change in the *orientation* of the body will be the same whichever order the motions are applied. It is not hard to see this intuitively.

This however is not so for the change in the position of the center of mass. Since the orientation does not change, the change in the height of the center of mass is the same whichever order the motions are applied. It is not hard to see intuitively (drawing arrows in the horizontal plane) that the change in the horizontal position is not the same when the order of two motions is changed. One may visual the difference between these two changes (draw another arrow, the difference), and ask what happens as the sizes of the motions approach zero, while maintaining their relative size. One may easily see that the *directions* in the results of lemma 9.4.1 are correct. It will appear below that for the flatness conditions to be satisfied, only the directions must be as in lemma 9.4.1.

Proposition 9.4.2. *The basis $\{\gamma_1, \gamma_2, \kappa\}$ satisfies the flatness conditions, equations 7.2 and 7.3 on page 51, which in this case become*

$$\begin{aligned} K(\gamma_1, [\gamma_1, \gamma_2]) &= 0, \\ K(\gamma_2, [\gamma_1, \gamma_2]) &= 0, \\ K(\kappa, [\kappa, \gamma_1]) &= 0 \qquad \text{and} \\ K(\kappa, [\kappa, \gamma_2]) &= 0. \end{aligned}$$

Proof. We calculate

$$
\begin{aligned}
K(\gamma_1, [\gamma_1, \gamma_2]) &= m(-rA(\epsilon_3)W\epsilon_3)^T\{rA(\epsilon_3)A(\epsilon_3)W\epsilon_3\} \\
&= 0. \\
K(\gamma_2, [\gamma_1, \gamma_2]) &= m(aA(\epsilon_3)W\epsilon_3)^T\{rA(\epsilon_3)A(\epsilon_3)W\epsilon_3\} \\
&= 0. \\
K(\kappa, [\kappa, \gamma_1]) &= m(-d\epsilon_3 + hW\epsilon_3)^T(-r'\sin\theta - r\cos\theta)A(\epsilon_3)W\epsilon_3 \\
&= 0 \qquad \text{and} \\
K(\kappa, [\kappa, \gamma_2]) &= m(-d\epsilon_3 + hW\epsilon_3)^T(a'\sin\theta - r)A(\epsilon_3)W\epsilon_3 \\
&= 0.
\end{aligned}
$$

□

By corollary 8.4.1 on page 68, we conclude that the reduced motion on Q/G is Hamiltonian.

9.5 The Adjoint Equation

In this subsection, we will find the form of the adjoint equation, equation 7.11 on page 57, or

$$\kappa(K(\gamma_i, \gamma_k)\Omega^k) = \{K([\kappa, \gamma_i], \gamma_k)) + K(\kappa, [\gamma_k, \gamma_i])\}\Omega^k,$$

for this problem.

We will however use suffixes for the components of Ω, in order to be consistent with the notation used in this section.

We will again work from

$$
\begin{aligned}
\gamma_1 W &= WA(\epsilon_3), \\
\gamma_1 x &= -rA(\epsilon_3)W\epsilon_3, \\
\gamma_2 W &= WA(W^T\epsilon_3) \\
&= A(\epsilon_3)W, \\
\gamma_2 x &= aA(\epsilon_3)W\epsilon_3, \\
\kappa W &= WA(A(W^T\epsilon_3)\epsilon_3) \\
&= A(A(\epsilon_3)W\epsilon_3)W \qquad \text{and} \\
\kappa x &= -d\epsilon_3 + hW\epsilon_3.
\end{aligned}
$$

together with

$$
\begin{aligned}
[\gamma_1, \gamma_2] &= (0, rA(\epsilon_3)A(\epsilon_3)W\epsilon_3), \\
[\kappa, \gamma_1] &= (0, -(r'\sin\theta + r\cos\theta)A(\epsilon_3)W\epsilon_3) \qquad \text{and} \\
[\kappa, \gamma_2] &= (0, (a'\sin\theta - r)A(\epsilon_3)W\epsilon_3).
\end{aligned}
$$

Proposition 9.5.1. *The adjoint equation for the axially symmetric rolling body becomes*

$$
\begin{aligned}
&\kappa\{(J_3 + mr^2\sin^2\theta)\Omega_1 + (J_3\cos\theta - mar\sin^2\theta)\Omega_2\} \\
&\qquad = mr\sin^2\theta(r'\sin\theta + r\cos\theta)\Omega_1 - m\sin^2\theta\{a(r'\sin\theta) - r^2\}\Omega_2 \\
&\qquad\quad \textit{and} \\
&\kappa\{(J_3\cos\theta - mar\sin^2\theta)\Omega_1 + (J_3\cos^2\theta + (J_1 + ma^2)\sin^2\theta)\Omega_2\} \\
&\qquad = -mr\sin^2\theta(a'\sin\theta + a\cos\theta)\Omega_1 + ma\sin^2\theta(a'\sin\theta - r)\Omega_2
\end{aligned}
$$

Proof. We calculate

$$
\begin{aligned}
K(\gamma_1,\gamma_1) &= \epsilon_3^T J\epsilon_3 + m(-rA(\epsilon_3)W\epsilon_3)^T(-rA(\epsilon_3)W\epsilon_3)\\
&= J_3 + mr^2\sin^2\theta.\\
K(\gamma_1,\gamma_2) &= \epsilon_3^T JW^{-T}\epsilon_3 + m(-rA(\epsilon_3)W\epsilon_3)^T(aA(\epsilon_3)W\epsilon_3)\\
&= J_3\cos\theta - mar\sin^2\theta.\\
K(\gamma_2,\gamma_2) &= \epsilon_3^T WJW^{-T}\epsilon_3 + m(aA(\epsilon_3)W\epsilon_3)^T(aA(\epsilon_3)W\epsilon_3)\\
&= \epsilon_3^T WJ\{(\epsilon_3^T W^{-T}\epsilon_3)\epsilon_3\} +\\
&\qquad \epsilon_3^T WJ\{W^{-T}\epsilon_3 - (\epsilon_3^T W^{-T}\epsilon_3)\epsilon_3\} + ma^2\sin^2\theta\\
&= J_3\cos^2\theta + (J_1 + ma^2)\sin^2\theta.\\
K([\kappa,\gamma_1],\gamma_1) &= m\{-(r'\sin\theta + r\cos\theta)A(\epsilon_3)W\epsilon_3\}^T(-rA(\epsilon_3)W\epsilon_3)\\
&= mr(r'\sin\theta + r\cos\theta)\sin^2\theta.\\
K([\kappa,\gamma_1],\gamma_2) &= m\{-(r'\sin\theta + r\cos\theta)A(\epsilon_3)W\epsilon_3\}^T(aA(\epsilon_3)W\epsilon_3)\\
&= -ma(r'\sin\theta + r\cos\theta)\sin^2\theta.\\
K([\kappa,\gamma_2],\gamma_1) &= m\{(a'\sin\theta - r)A(\epsilon_3)W\epsilon_3\}^T(-rA(\epsilon_3)W\epsilon_3)\\
&= -mr(a'\sin\theta - r)\sin^2\theta.\\
K([\kappa,\gamma_2],\gamma_2) &= m\{(a'\sin\theta - r)A(\epsilon_3)W\epsilon_3\}^T(aA(\epsilon_3)W\epsilon_3)\\
&= ma(a'\sin\theta - r)\sin^2\theta\\
&\text{and}\\
K(\kappa,[\gamma_1,\gamma_2]) &= m(-d\epsilon_3 + hW\epsilon_3)^T(rA(\epsilon_3)A(\epsilon_3)W\epsilon_3)\\
&= -mhr\sin^2\theta\\
&= -mr(r + a\cos\theta)\sin^2\theta.
\end{aligned}
$$

□

We now observe that since θ is group invariant, it induces a coordinate on Q/G. Recalling $\kappa\theta = \sin\theta$ we have

Corollary 9.5.1. *Using θ as a coordinate for Q/G, the adjoint equation* *pushes down to*

$$(9.10)\quad \frac{d}{d\theta}\{(J_3 + mr^2\sin^2\theta)\Omega_1 + (J_3\cos\theta - mar\sin^2\theta)\Omega_2\}$$
$$= mr\sin\theta\left(\frac{dr}{d\theta}\sin\theta + r\cos\theta\right)\Omega_1$$
$$- m\sin\theta\left(a\frac{dr}{d\theta}\sin\theta - r^2\right)\Omega_2$$

and

$$(9.11)\quad \frac{d}{d\theta}\{(J_3\cos\theta - mar\sin^2\theta)\Omega_1 + [J_3\cos^2\theta + (J_1 + ma^2)\sin^2\theta]\Omega_2\}$$
$$= -mr\sin\theta\left(\frac{da}{d\theta}\sin\theta + a\cos\theta\right)\Omega_1$$
$$+ ma\sin\theta\left(\frac{da}{d\theta}\sin\theta - r\right)\Omega_2$$

on Q/G.

In order to compare our result with various traditional results, it will be helpful to be able to relate the quasi-velocities of our theory with the angular velocities.

From the general theory, we know that the tangent vector to the trajectory, $\tau(t)$, will be a linear combination of γ_1, γ_2 and κ. So we have

$$\begin{aligned}
\omega_1(\Omega_1\gamma_1 + \Omega_2\gamma_2 + v\kappa) &= \Omega_1\omega_1(\gamma_1) + \Omega_2\omega_1(\gamma_2) + v\omega_1(\kappa)\\
&= \Omega_2(\epsilon_1^T W^{-T}\epsilon_3) + v\{\epsilon_1^T A(W^{-T}\epsilon_3)\epsilon_3\}\\
&= \Omega_2(\epsilon_1^T W^{-T}\epsilon_3) + v(\epsilon_2^T W^{-T}\epsilon_3)\\
&= \Omega_2(\epsilon_3^T W^{-}\epsilon_1) + v(\epsilon_3^T W^{-}\epsilon_2).\\
\omega_2(\Omega_1\gamma_1 + \Omega_2\gamma_2 + v\kappa) &= \Omega_1\omega_2(\gamma_1) + \Omega_2\omega_2(\gamma_2) + v\omega_2(\kappa)\\
&= \Omega_2(\epsilon_2^T W^{-T}\epsilon_3) + v\{\epsilon_2^T A(W^{-T}\epsilon_3)\epsilon_3\}\\
&= \Omega_2(\epsilon_2^T W^{-T}\epsilon_3) - v(\epsilon_1^T W^{-T}\epsilon_3)\\
&= \Omega_2(\epsilon_3^T W^{-}\epsilon_2) - v(\epsilon_3^T W^{-}\epsilon_1) \qquad \text{and}\\
\omega_3(\Omega_1\gamma_1 + \Omega_2\gamma_2 + v\kappa) &= \Omega_1\omega_3(\gamma_1) + \Omega_2\omega_3(\gamma_2) + v\omega_3(\kappa)\\
&= \Omega_1 + \Omega_2(\epsilon_3^T W^{-T}\epsilon_3)\\
&= \Omega_1 + \cos\theta\,\Omega_2.
\end{aligned}$$

From the first two equations we have

$$\begin{aligned}\{\epsilon_3^T W\epsilon_1\}\omega_1(\tau)+\{\epsilon_3^T W\epsilon_2\}\omega_2(\tau) &= \Omega_2\{(\epsilon_3^T W\epsilon_1)^2+(\epsilon_3^T W\epsilon_2)^2\}\\ &= \Omega_2\{1-(\epsilon_3^T W\epsilon_3)^2\}\\ &= \Omega_2\{1-\cos^2\theta\}\\ &= \sin^2\theta\,\Omega_2.\end{aligned}$$

Lemma 9.5.1. *The relations*

$$\begin{aligned}\{\sin^2\theta(t)\}\Omega_2\circ\theta(t) &= \{\epsilon_3^T W(t)\epsilon_1\}\omega_1\circ\tau(t)+\{\epsilon_3^T W(t)\epsilon_2\}\omega_2\circ\tau(t)\\ &\textit{and}\\ \varpi_3\circ\theta(t) &= \omega_3\circ\tau(t)\end{aligned}$$

hold, where ϖ_3 *is defined by*

$$\varpi_3=\Omega_1+\cos\theta\,\Omega_2.$$

Corollary 9.5.2. *Equations 9.10 and 9.11 may be rewritten using* ϖ_3 *and* Ω_2 *in the form*

$$a\frac{d}{d\theta}\{J_3\varpi_3\}+r\frac{d}{d\theta}\{J_3\cos\theta\,\varpi_3+J_1\sin^2\theta\,\Omega_2\}=0 \tag{9.12}$$

and

$$\frac{d}{d\theta}\{J_3\varpi_3\}+mr\sin\theta\frac{d}{d\theta}\{\sin\theta[r\varpi_3-(a+r\cos\theta)\Omega_2]\} = mr\sin\theta(r+a\cos\theta)\Omega_2. \tag{9.13}$$

Proof. Equation 9.12 may be obtained as a times equation 9.10 plus r times equation 9.11. In the first step we obtain

$$\begin{aligned}&a\frac{d}{d\theta}\{J_3\varpi_3\}+ma\frac{d}{d\theta}\{r\sin^2\theta(r\Omega_1-a\Omega_2)\}\\ &\quad+r\frac{d}{d\theta}\{J_3\cos\theta\,\varpi_3+J_1\sin^2\theta\,\Omega_2\}-mr\frac{d}{d\theta}\{a\sin^2\theta(r\Omega_1-a\Omega_2)\}\\ &\qquad=m\sin^2\theta\left\{a\frac{dr}{d\theta}-r\frac{da}{d\theta}\right\}(r\Omega_1-a\Omega_2).\end{aligned}$$

which simplifies to the desired result.

Equation 9.13 is a rearrangement of equation 9.10. In the first step equation 9.10 becomes

$$\frac{d}{d\theta}\{J_3\varpi_3\} + m\frac{d}{d\theta}\{r\sin^2\theta(r\Omega_1 - a\Omega_2)\}$$
$$= m\sin^2\theta\frac{dr}{d\theta}(r\Omega_1 - a\Omega_2) + mr^2\sin\theta(\cos\theta\,\Omega_1 + \Omega_2).$$

followed by

$$\frac{d}{d\theta}\{J_3\varpi_3\} + mr\frac{d}{d\theta}\{\sin^2\theta(r\Omega_1 - a\Omega_2)\} = mr^2\sin\theta(\cos\theta\,\Omega_1 + \Omega_2).$$

and then

$$\frac{d}{d\theta}\{J_3\varpi_3\} + mr\sin\theta\frac{d}{d\theta}\{\sin\theta(r\Omega_1 - a\Omega_2)\} = mr\sin\theta(r + a\cos\theta)\Omega_2.$$

which simplifies to the desired result. □

Equations 9.12 and 9.13 are obtained in Gallop [16, page 371] using force arguments. The variables used there differ slightly from those used here. Equivalent equations are found in Routh [31, Article 241a].

Example. The Rolling Disk

For the rolling disk, we have

$$\begin{aligned} a(\theta) &= -\rho\frac{\cos\theta}{\sin\theta}, \\ r(\theta) &= \rho\frac{1}{\sin\theta}, \\ \frac{da}{d\theta} &= \rho\frac{1}{\sin^2\theta} \qquad \text{and} \\ \frac{dr}{d\theta} &= -\rho\frac{\cos\theta}{\sin^2\theta}. \end{aligned}$$

From this we obtain

$$\begin{aligned} a + r\cos\theta &= 0, \\ r + a\cos\theta &= \rho\sin\theta \qquad \text{and} \\ r\sin\theta &= \rho. \end{aligned}$$

Equations 9.12 and 9.13 become

$$\begin{aligned} \frac{d}{d\theta}\{J_1 \sin^2\theta\, \Omega_2\} &= J_3 \sin\theta\, \varpi_3 \qquad\qquad \text{and} \\ (J_3 + m\rho^2)\frac{d\varpi_3}{d\theta} &= m\rho^2 \sin\theta\, \Omega_2. \end{aligned}$$

respectively. For ϖ_3 we have

$$J_1(J_3 + m\rho^2)\frac{d}{d\theta}\left\{\sin\theta \frac{d\varpi_3}{d\theta}\right\} = m\rho^2 J_3 \sin\theta\, \varpi_3.$$

If instead of θ we use $p = \cos\theta$ as a coordinate for Q/G. this becomes

$$\frac{d}{dp}\left\{(1-p^2)\frac{d\varpi_3}{dp}\right\} - \frac{J_3 m\rho^2}{J_1(J_3 + m\rho^2)}\varpi_3 = 0.$$

This equation is obtained in Pars [29. Section 8.12] using Lagrange's Equations. and in Routh [31. Article 244a] using force arguments. It is an equation of Legendre type. Gallop [16] observed that an equation of Legendre type resulted in this case. but did not do the calculation. □

Example. Routh's Model for a Top

For Routh's model of a top. r and a are constant. and positive. We will discuss the $r = 0$ case seperately below.

In this case. equation 9.12 immediately yields

$$(a + r\cos\theta)J_3\varpi_3 + r\sin^2\theta\, J_1\Omega_2 = C_1.$$

where C_1 is a constant. This constant is found in Routh [31. Article 243] using force arguments. It is shown there that this result will also hold when the top is subject to slipping friction at the point of contact. The constant C_1 is usually called Jellett's integral. the functional expression for this having been introduced by him [22]. although he was unaware that it was a true constant of the motion. Jellett's integral was used by Gallop ([16]) and Ebenfeld and Scheck ([13]). to investigate the "rising" behaviour (and. in the case of [13]. other asymptotic behaviour) of the conventional spinning top and of the "tippe top". in the presence of dissipative sliding friction. Both analyses. at their heart. determine the configuration for which the total energy is minimized. subject to the existence of Jellett's integral.

Equation 9.13 becomes

$$
\begin{aligned}
(J_3 + mr^2\sin^2\theta)\frac{d\varpi_3}{d\theta} + mr^2\sin\theta\cos\theta\,\varpi_3 \\
- mr\sin\theta\cos\theta(a + r\cos\theta)\Omega_2 + mr^2\sin^3\theta\,\Omega_2 \\
- mr\sin^2\theta(a + r\cos\theta)\frac{d\Omega_2}{d\theta} \\
= mr\sin\theta(r + a\cos\theta)\Omega_2.
\end{aligned}
$$

then

$$
\begin{aligned}
(J_3 + mr^2\sin^2\theta)\frac{d\varpi_3}{d\theta} + mr^2\sin\theta\cos\theta\,\varpi_3 \\
- mr\sin^2\theta(a + r\cos\theta)\frac{d\Omega_2}{d\theta} \\
- 2mr\sin\theta\cos\theta(a + r\cos\theta)\Omega_2 = 0
\end{aligned}
$$

and finally

$$
\begin{aligned}
(J_3 + mr^2\sin^2\theta)\frac{d\varpi_3}{d\theta} + mr^2\sin\theta\cos\theta\,\varpi_3 \\
- mr(a + r\cos\theta)\frac{d}{d\theta}\{\sin^2\theta\,\Omega_2\} = 0.
\end{aligned}
$$

Equation 9.12 may now be used to eliminate Ω_2. We obtain

$$
\begin{aligned}
J_1(J_3 + mr^2\sin^2\theta)\frac{d\varpi_3}{d\theta} + J_1 mr^2\sin\theta\cos\theta\,\varpi_3 \\
+ m(a + r\cos\theta)J_3\frac{d}{d\theta}\{(a + r\cos\theta)\varpi_3\} = 0.
\end{aligned}
$$

Multiplying this by $2\varpi_3$ results in a perfect differential. We obtain

$$
\frac{d}{d\theta}\{[J_1(J_3 + mr^2\sin^2\theta) + mJ_3(a + r\cos\theta)^2]\varpi_3^2\} = 0.
$$

This result is found in Routh [31, Article 243] using force arguments. The solution to this is

$$
\varpi_3 = \frac{C_2}{\sqrt{J_1(J_3 + mr^2\sin^2\theta) + mJ_3(a + r\cos\theta)^2}}.
$$

where C_2 is a constant. In particular we notice that ϖ_3 will never be zero, unless it has zero as its constant value.

We may now solve for Ω_2, obtaining

$$\Omega_2 = \frac{1}{rJ_1\sin^2\theta}\left\{C_1 - \frac{C_2J_3(a+r\cos\theta)}{\sqrt{J_1(J_3+mr^2\sin^2\theta)+mJ_3(a+r\cos\theta)^2}}\right\}.$$

We will use this result in the following subsection to obtain an expression for the potential energy in the reduced problem.

The constant C_1 (Jellett's integral) has an interpretation in terms of the angular momentum of the top. We ask the reader to recall equation 2.10 on page 18, and the notation in use in that subsection. There, the angular momentum about the origin was the quantity $P = mA(x)\dot{x} + WJ\omega$. The expression $WJ\omega$ was called the angular momentum of the body. Also recall lemma 9.5.1. Then we have

$$w_3^T\{WJ\omega\} = (We_3)^TWJ\omega = J_3\omega_3$$

and

$$\begin{aligned} e_3^T\{WJ\omega\} &= e_3^TW\{J_1(\omega_1e_1+\omega_2e_2)+J_3\omega_3e_3\} \\ &= J_3\cos\theta\,\omega_3 + J_1\{(e_3^TWe_1)\omega_1 + (e_3^TWe_2)\omega_2\} \\ &= J_3\cos\theta\,\omega_3 + J_1\sin^2\theta\,\Omega_2. \end{aligned}$$

So we have $C_1 = a(w_3^TWJ\omega) + r(e_3^TWJ\omega)$, and C_1 is a weighted sum of components along the vertical axis and the symmetry axis, of the angular momentum of the body.

Also, if we let σ be the vector field given in $(\omega, \dot{x})$-space by

$$\sigma = a\gamma_1 + r\gamma_2 = (ae_3 + rW^Te_3, 0),$$

then σ satisfies the conditions on σ given in proposition 5.1.1 on page 40, and so generates a constant of the motion. This constant is in fact C_1.

To see this, we have $\sigma \in D$ by design, and the flow of σ is given by

$$\sigma_s(W, x) = (\overline{W}, \overline{x})$$

with $\overline{W} = R_3(rs)WR_3(as)$ and $\overline{x} = x$. Differentiating this gives $\overline{\omega} = R_3(as)^T\omega$ and $\dot{\overline{x}} = \dot{x}$, in the usual way, so that σ fully preserves the Lagrangian.

$$L = \frac{1}{2}\omega^TJ\omega + \frac{1}{2}m\dot{x}^T\dot{x} - mge_3^Tx.$$

If $\tau(t)$ is the tangent vector to the trajectory, the constant of the motion generated by σ then must be

$$K(\tau,\sigma) = \omega(\tau)^T J\{a\epsilon_3 + rW^T\epsilon_3\} = C_1.$$

□

Example. Lagrange Top

For the case of the Lagrange top with $r = 0$ and a fixed, Jellett's integral simply requires that ϖ_3 should be a constant. This also follows from putting $r = 0$ in the expression for ϖ_3. However, we are unable to solve for Ω_2 using equations 9.12 and 9.13, for $r = 0$.

To obtain ϖ_3 and Ω_2 in the case of the Lagrange top, we return to equations 9.10 and 9.11, which for $r = 0$ become

$$\begin{aligned} J_3\frac{d\varpi_3}{d\theta} &= 0 \qquad \text{and} \\ \frac{d}{d\theta}\{J_3\cos\theta\,\varpi_3 + (J_1 + ma^2)\sin^2\theta\,\Omega_2\} &= 0. \end{aligned}$$

or

$$\begin{aligned} \varpi_3 &= C_2 \qquad \text{and} \\ J_3\cos\theta\,\varpi_3 + (J_1 + ma^2)\sin^2\theta\,\Omega_2 &= C_1. \end{aligned}$$

We may write the solution in the form

$$\begin{aligned} \varpi_3 &= C_2 \qquad \text{and} \\ \Omega_2 &= \frac{1}{(J_1 + ma^2)\sin^2\theta}\{C_1 - C_2J_3\cos\theta\}. \end{aligned}$$

The constants C_1 and C_2 have an interpretation in terms of the angular momentum of the top about the point of contact of the top with the ground. We ask the reader to recall equation 2.10 on page 18, and the notation in use in that subsection. There, the angular momentum about the origin was the quantity $P = mA(x)\dot{x} + WJ\omega$. The expression $WJ\omega$ was called the angular momentum of the body. We take the origin to be at the point of contact, so that $x = aw_3 = aWe_3$ holds. Also recall lemma 9.5.1. Then we have

$$w_3^TP = mw_3^TA(x)\dot{x} + w_3^T(WJ\omega) = (We_3)^TWJ\omega = J_3\omega_3 = J_3C_2.$$

Thus $J_3 C_2$ is the component of angular momentum along the axis of symmetry. We also have

$$\begin{aligned}
\epsilon_3^T P = ma^2 \epsilon_3^T A(W\epsilon_3) W A(\omega)\epsilon_3 + \epsilon_3^T W J\omega = \\
- ma^2 \epsilon_3^T W A(\epsilon_3) A(\epsilon_3)\omega + \epsilon_3^T W J\omega = \\
- ma^2 \cos\theta\, \omega_3 + \epsilon_3^T W (J + ma^2 I)\omega = \\
J_3 \cos\theta\, \omega_3 + (J_1 + ma^2)\{(\epsilon_3^T W \epsilon_1)\omega_1 + (\epsilon_3^T W \epsilon_2)\omega_2\} = \\
J_3 \cos\theta\, \omega_3 + (J_1 + ma^2)\sin^2\theta\, \Omega_2 = \\
C_1.
\end{aligned}$$

Thus C_1 is the component of angular momentum along the vertical axis.

Since the Lagrange top is a limiting case of the Routh top, it is possible to obtain this solution from the Routh top solution, by letting r approach zero, while adjusting the constants of integration. One way of choosing the constants is to put

$$\begin{aligned}
C_1 &= \frac{rJ_1}{J_1 + ma^2}\overline{C}_1 + aJ_3\overline{C}_2 \\
&\text{and} \\
C_2 &= \sqrt{J_1 J_3 + J_1 m r^2 + J_3 m a^2}\,\overline{C}_2
\end{aligned}$$

into the Routh top solution, and take the limit. The calculation is not difficult, but also not very instructive. The quantities $\overline{C}_1$ and $\overline{C}_2$ become the constants in the Lagrange top solution above. This choice for C_1 and C_2 is the one that ensures that

$$\varpi_3\left(\frac{\pi}{2}\right) \quad \text{and} \quad \Omega_2\left(\frac{\pi}{2}\right)$$

have the same value for each value of r. □

9.6 The Reduced Kinetic Energy and Potential Function

In this subsection, we will find the form of the kinetic energy and the potential function for the reduced problem.

Recall (lemma 8.4.1) that the reduced potential function $\overline{U}$ is determined by

$$\overline{U} \circ \Lambda = U + \frac{1}{2}K(\Omega, \Omega),$$

while the reduced kinetic energy inner product $\overline{K}$ is induced by the requirement (lemma 8.4.2)

$$\overline{K}(T\Lambda(\tau), T\Lambda(\tau)) = K(\tau, \tau), \forall \tau \in H \div R.$$

For the axially symmetric rolling body the potential function may be written as

$$U = mg(r + a\cos\theta).$$

The kinetic energy inner product once again is given by

$$K((\lambda, u), (\mu, v)) = \lambda^T J\mu + m u^T v.$$

We also have

$$\begin{aligned} \Omega &= (\varpi_3 - \cos\theta\, \Omega_2)\gamma_1 + \Omega_2\gamma_2 \\ &= \varpi_3\gamma_1 + \Omega_2(-\cos\theta\, \gamma_1 + \gamma_2), \end{aligned}$$

and the elements of our aligned basis are

$$\begin{aligned} \gamma_1 &= (\epsilon_3, -rA(\epsilon_3)W\epsilon_3), \\ \gamma_2 &= (W^T\epsilon_3, aA(\epsilon_3)W\epsilon_3) \qquad \text{and} \\ \kappa &= (A(W^T\epsilon_3)\epsilon_3, -d\epsilon_3 + hW\epsilon_3) \end{aligned}$$

First we will examine the kinetic energy. We have earlier seen that, for the axially symmetric rolling body,

$$\kappa\theta = \sin\theta$$

holds, when κ is thought of as a differentiation operator. This means that κ is the horizontal lift of $\sin\theta\, \partial/\partial\theta$. Therefore we have

$$\overline{K}\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}\right) = \frac{1}{\sin^2\theta} K(\kappa, \kappa).$$

with

$$\begin{aligned}\kappa &= (A(W^{-T}e_3)e_3, -(a + r\cos\theta)e_3 + (r + a\cos\theta)W^{-}e_3)\\ &= (A(W^{-T}e_3)e_3, a(-e_3 + \cos\theta\, W^{-}e_3) + r(-\cos\theta\, e_3 + W^{-}e_3)).\end{aligned}$$

Observe that $A(W^{-T}e_3)e_3$ is a vector of magnitude $\sin\theta$, orthogonal to e_3, and so an eigenvector of J with eigenvalue J_1. We now calculate

$$\begin{aligned}\overline{K}\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}\right) &= \frac{1}{\sin^2\theta}K(\kappa,\kappa)\\ &= \frac{1}{\sin^2\theta}\{J_1\sin^2\theta + m[a^2(1-\cos^2\theta)\\ &\qquad + 2ar(\cos\theta - \cos^3\theta) + r^2(-\cos^2\theta + 1)]\}\\ &= J_1 + m(a^2 + 2ar\cos\theta + r^2).\end{aligned}$$

To examine the reduced potential function, we need an actual solution to the adjoint equation. So we do this for examples.

Example. Routh's Model for a Top

Our solutions in this case were

$$\overline{\omega}_3 = \frac{C_2}{\sqrt{J_1(J_3 + mr^2\sin^2\theta) + mJ_3(a + r\cos\theta)^2}}$$

and

$$\Omega_2 = \frac{1}{rJ_1\sin^2\theta}\left\{C_1 - \frac{C_2J_3(a + r\cos\theta)}{\sqrt{J_1(J_3 + mr^2\sin^2\theta) + mJ_3(a + r\cos\theta)^2}}\right\}.$$

where C_1 and C_2 are constants.

In order to simplify the calculations, we note that $\overline{U}$ will necessarily be of the form

$$\overline{U}(\theta) = mg(r + a\cos\theta) + \frac{1}{2r^2J_1^2\sin^2\theta}\{C_1^2P + 2C_1C_2Q + C_2^2R\},$$

and proceed to find the coefficients P, Q and R. If we put

$$\beta = \frac{1}{\sqrt{J_1(J_3 + mr^2\sin^2\theta) + mJ_3(a + r\cos\theta)^2}},$$

then we have

$$\Omega = \frac{1}{J_1 r\sin^2\theta}\{C_1\overline{\gamma}_1 + C_2\beta\overline{\gamma}_2\}.$$

with

$$
\begin{aligned}
\bar{\gamma}_1 &= -\cos\theta\,\gamma_1 + \gamma_2 \\
&= (-\cos\theta\,\epsilon_3 + W^T\epsilon_3.(a + r\cos\theta)A(\epsilon_3)W\epsilon_3) \\
&\text{and} \\
\bar{\gamma}_2 &= J_1 r \sin^2\theta\,\gamma_1 - J_3(a + r\cos\theta)\bar{\gamma}_1 \\
&= (J_1 r \sin^2\theta\,\epsilon_3 + J_3(a + r\cos\theta)(\cos\theta\,\epsilon_3 - W^T\epsilon_3). \\
&\qquad - [J_1 r^2 \sin^2\theta + J_3(a + r\cos\theta)^2]A(\epsilon_3)W\epsilon_3).
\end{aligned}
$$

Observe that both $-\cos\theta\,\epsilon_3 + W^T\epsilon_3$ and $A(\epsilon_3)W\epsilon_3$ are vectors of magnitude $\sin\theta$. They are also both orthogonal to ϵ_3, and so are eigenvectors of J with eigenvalue J_1. We may now calculate

$$
\begin{aligned}
P &= \frac{1}{\sin^2\theta}K(\bar{\gamma}_1, \bar{\gamma}_1) \\
&= \frac{1}{\sin^2\theta}\{J_1 \sin^2\theta + m\sin^2\theta(a + r\cos\theta)^2\} \\
&= J_1 + m(a + r\cos\theta)^2. \\
Q &= \beta\frac{1}{\sin^2\theta}K(\bar{\gamma}_1, \bar{\gamma}_2) \\
&= \beta\frac{1}{\sin^2\theta}\{-J_1 J_3(a + r\cos\theta)\sin^2\theta \\
&\qquad - m\sin^2\theta(a + r\cos\theta)[J_1 r^2 \sin^2\theta + J_3(a + r\cos\theta)^2]\} \\
&= -\beta(a + r\cos\theta)\{J_1(J_3 + mr^2\sin^2\theta) + mJ_3(a + r\cos\theta)^2\} \\
&= -(a + r\cos\theta)\sqrt{J_1(J_3 + mr^2\sin^2\theta) + mJ_3(a + r\cos\theta)^2} \\
&\text{and} \\
R &= \beta^2\frac{1}{\sin^2\theta}K(\bar{\gamma}_2, \bar{\gamma}_2) \\
&= \beta^2\frac{1}{\sin^2\theta}\{J_3 J_1^2 r^2 \sin^4\theta + J_1 J_3^2 \sin^2\theta(a + r\cos\theta)^2 \\
&\qquad + m\sin^2\theta[J_1 r^2\sin^2\theta + J_3(a + r\cos\theta)^2]^2\} \\
&= \beta^2[J_1 r^2 \sin^2\theta + J_3(a + r\cos\theta)^2] \\
&\qquad \{J_1 J_3 + m[J_1 r^2 \sin^2\theta + J_3(a + r\cos\theta)^2]\} \\
&= J_1 r^2 \sin^2\theta + J_3(a + r\cos\theta)^2.
\end{aligned}
$$

Finally then, the reduced potential function is

$$\begin{aligned}
\bar{V}(\theta) &= mg(r + a\cos\theta) + \frac{1}{2r^2J_1^2\sin^2\theta}\{ \\
&\quad C_1^2[J_1 + m(a + r\cos\theta)^2] \\
&\quad - 2C_1C_2(a + r\cos\theta) \\
&\quad \sqrt{J_1(J_3 + mr^2\sin^2\theta) + mJ_3(a + r\cos\theta)^2} \\
&\quad + C_2^2[J_1r^2\sin^2\theta + J_3(a + r\cos\theta)^2] \} \\
&= \frac{1}{2J_1}\left(C_2^2 - \frac{mC_1^2}{J_3}\right) + mg(r + a\cos\theta) \\
&\quad + \frac{1}{2r^2J_1^2J_3\sin^2\theta}\{ \\
&\quad C_1\sqrt{J_1(J_3 + mr^2\sin^2\theta) + mJ_3(a + r\cos\theta)^2} \\
&\quad - C_2J_3(a + r\cos\theta)\}^2.
\end{aligned}$$

For the total reduced energy we have

$$\begin{aligned}
E &= \frac{1}{2}\{J_1 + m(a^2 + 2ar\cos\theta + r^2)\}\dot{\theta}^2 \\
&\quad + \frac{1}{2J_1}\left(C_2^2 - \frac{mC_1^2}{J_3}\right) + mg(r + a\cos\theta) \\
&\quad + \frac{1}{2r^2J_1^2J_3\sin^2\theta}\{ \\
&\quad C_1\sqrt{J_1(J_3 + mr^2\sin^2\theta) + mJ_3(a + r\cos\theta)^2} \\
&\quad - C_2J_3(a + r\cos\theta)\}^2.
\end{aligned}$$

This is effectively the same expression briefly examined in Routh [31, Article 243a]. We will analyse this expression briefly in the next subsection. □

Example. Lagrange Top

Our solutions in this case were

$$\pi_3 = C_2$$

and

$$\Omega_2 = \frac{1}{(J_1 + ma^2)\sin^2\theta}\{C_1 - C_2J_3\cos\theta\}.$$

where C_1 and C_2 are constants.

We note that $\overline{V}$ will be of the form

$$\overline{V}(\theta) = mga\cos\theta + \frac{1}{2(J_1 + ma^2)^2\sin^2\theta}\{C_1^2 P + 2C_1C_2Q + C_2^2 R\}.$$

and proceed to find the coefficients P, Q and R. We have

$$\Omega = \frac{1}{(J_1 + ma^2)\sin^2\theta}\{C_1\overline{\gamma}_1 + C_2\overline{\gamma}_2\}.$$

with

$$\begin{aligned}
\overline{\gamma}_1 &= -\cos\theta\,\gamma_1 + \gamma_2\\
&= (-\cos\theta\,\epsilon_3 + W^T\epsilon_3,\ aA(\epsilon_3)W\epsilon_3)\\
&\text{and}\\
\overline{\gamma}_2 &= (J_1 + ma^2)\sin^2\theta\,\gamma_1 - J_3\cos\theta\overline{\gamma}_1\\
&= ((J_1 + ma^2)\sin^2\theta\,\epsilon_3 + J_3\cos\theta(\cos\theta\,\epsilon_3 - W^T\epsilon_3),\\
&\qquad - J_3 a\cos\theta\, A(\epsilon_3)W\epsilon_3).
\end{aligned}$$

We may now calculate

$$\begin{aligned}
P &= \frac{1}{\sin^2\theta}K(\overline{\gamma}_1,\overline{\gamma}_1)\\
&= \frac{1}{\sin^2\theta}\{J_1\sin^2\theta + ma^2\sin^2\theta\}\\
&= J_1 + ma^2.\\
Q &= \frac{1}{\sin^2\theta}K(\overline{\gamma}_1,\overline{\gamma}_2)\\
&= \frac{1}{\sin^2\theta}\{-J_1J_3\cos\theta\sin^2\theta - mJ_3a^2\cos\theta\sin^2\theta\}\\
&= -J_3\cos\theta(J_1 + ma^2)\\
&\text{and}\\
R &= \frac{1}{\sin^2\theta}K(\overline{\gamma}_2,\overline{\gamma}_2)\\
&= \frac{1}{\sin^2\theta}\{J_3(J_1 + ma^2)^2\sin^4\theta + J_1J_3^2\cos^2\theta\sin^2\theta\\
&\qquad + mJ_3^2a^2\cos^2\theta\sin^2\theta\}\\
&= (J_1 + ma^2)\{J_3\sin^2\theta(J_1 + ma^2) + J_3^2\cos^2\theta\}.
\end{aligned}$$

The reduced potential function is

$$\begin{aligned}
\overline{U}(\theta) &= mga\cos\theta + \frac{1}{2(J_1 + ma^2)\sin^2\theta}\{ \\
&\qquad C_1^2 - 2C_1C_2J_3\cos\theta \\
&\qquad\quad + C_2^2[J_3\sin^2\theta(J_1 + ma^2) + J_3^2\cos^2\theta]\} \\
&= \frac{1}{2}C_2^2J_3 + mga\cos\theta + \frac{1}{2(J_1 + ma^2)\sin^2\theta}(C_1 - C_2J_3\cos\theta)^2.
\end{aligned}$$

For the total reduced energy we have

$$\begin{aligned}
E &= \frac{1}{2}(J_1 + ma^2)\dot{\theta}^2 + \frac{1}{2}C_2^2J_3 + mga\cos\theta \\
&\qquad + \frac{1}{2(J_1 + ma^2)\sin^2\theta}(C_1 - C_2J_3\cos\theta)^2.
\end{aligned}$$

This result may be compared for example with Arnold [2. Page 152]. □

9.7 Further Discussion of the Routh Top

In this subsection, we will give a brief examination of the reduced system for Routh's model of a top. This, once again, was the case of a spherical base.

In section 9.6, we showed that the kinetic energy of the reduced system is given by

$$K(\theta,\dot\theta) = \frac{1}{2}\{J_1 + m(a^2 + 2ar\cos\theta + r^2)\}\dot\theta^2,$$

and the reduced potential energy by

$$\begin{aligned} U(\theta) &= \frac{1}{2J_1}\left(C_2^2 - \frac{mC_1^2}{J_3}\right) + mg(r + a\cos\theta) \\ &\quad + \frac{1}{2r^2J_1^2J_3\sin^2\theta}\{ \\ &\quad C_1\sqrt{J_1(J_3 + mr^2\sin^2\theta) + mJ_3(a + r\cos\theta)^2} \\ &\quad - C_2J_3(a + r\cos\theta)\}^2. \end{aligned}$$

If we define functions $P, S, Q : (-1, 1) \to \Re$ by

$$P(z) = J_1 + m(a^2 + r^2 + 2arz),$$

$$S(z) =$$

$$C_1\sqrt{J_1(J_3 + mr^2(1 - z^2)) + mJ_3(a + rz)^2} - C_2J_3(a + rz)$$

and

$$Q(z) = mgaz + \frac{S(z)^2}{2r^2J_1^2J_3(1 - z^2)},$$

this becomes

$$\begin{aligned} K(\theta,\dot\theta) &= \frac{1}{2}P(\cos\theta)\dot\theta^2 \qquad \text{and} \\ U(\theta) &= \frac{1}{2J_1}\left(C_2^2 - \frac{mC_1^2}{J_3}\right) + mgr + Q(\cos\theta). \end{aligned}$$

Since P is positive on $[-1, 1]$, the qualitative behaviour of the reduced system may be determined from the potential function, as a point sliding on a potential curve. The constant term may be adjusted to any value by choice of the total energy. So the shape of the function U will determine the possible reduced motions.

What we need to know are

- the limits of U as $\theta \to 0, \pi$ and
- the local minima, maxima and other critical points of U

Case 1: $S(-1) = S(+1) = 0$

Observe that

$$\begin{aligned} S(-1) &= C_1\sqrt{J_1J_3 + mJ_3(a-r)^2} - C_2J_3(a-r) \qquad \text{and} \\ S(+1) &= C_1\sqrt{J_1J_3 + mJ_3(a+r)^2} - C_2J_3(a+r) \end{aligned}$$

hold.

It is not hard to see that since we have $r \neq 0$, this case requires $C_1 = C_2 = 0$. This means there is no rotation about either the symmetry axis or the vertical axis. The top either approaches the top position but reverses before reaching it, takes an infinite time to reach it, or passes through it, depending upon the value for the energy.

Case 2: $S(-1) \neq 0$ and $S(+1) \neq 0$

We have $Q(z) \to \infty$ as $z \to \pm 1$, or $U(\theta) \to \infty$ as $\theta \to 0, \pi$. In this case, U must have a local minimum, which corresponds to a motion of the top with a fixed value for θ. There will also be periodic motions about this value of θ.

If U has no other critical points, then all other motions are periodic. If on the other hand U does have other critical points, there are orbits which approach a limit point asymptotically. It is clear that the generic behaviour will be periodic motion.

Figure 9.2 shows a plot of the function Q produced by the Maple mathematical software.

Case 3: $S(-1) \neq 0$ and $S(+1) = 0$

S is analytic on an open interval containing $[-1, +1]$. So the expression

$$\frac{S(z)}{1-z}$$

is analytic at $z = +1$, and therefore so is Q. This in turn shows that U is analytic on the interval $(-\pi, \pi)$. We must have $U(\theta) \to \infty$ as $\theta \to \pm\pi$.

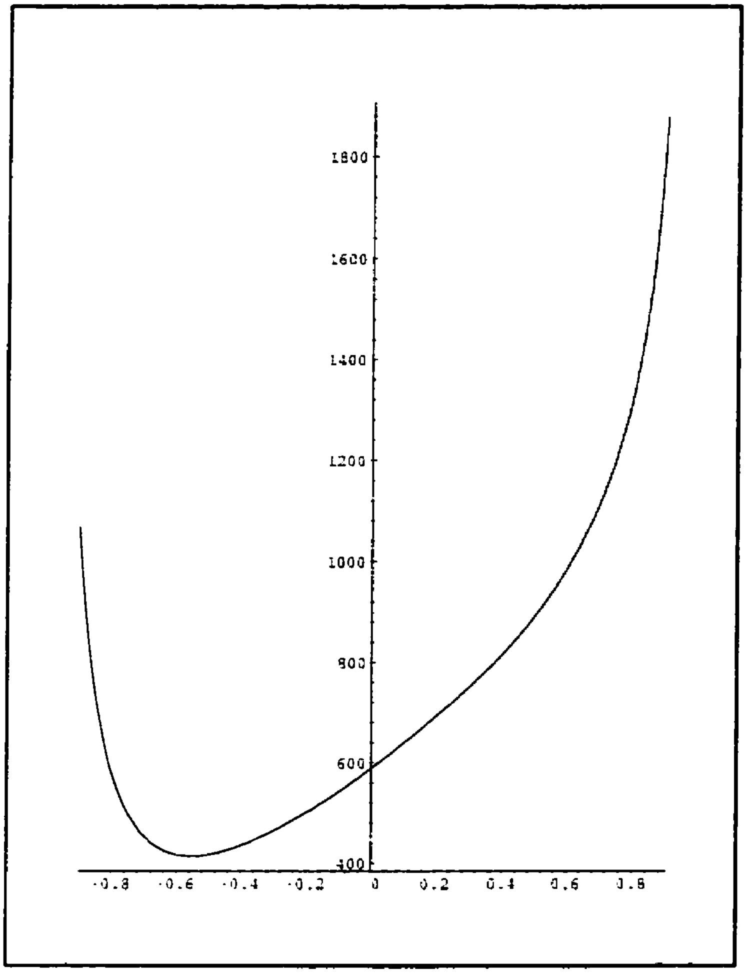


Figure 9.2: Plot of the Function Q

But then from

$$U'(\theta) = -Q'(\cos\theta)\sin\theta$$

we see that $U'(0) = 0$ holds, and from

$$U''(\theta) = Q''(\cos\theta)\sin^2\theta - Q'(\cos\theta)\cos\theta$$

that $U''(0) = -Q'(+1)$ holds.

To evaluate this, first note that

$$\frac{S(z)}{1-z^2} \to -\frac{S'(+1)}{2}$$

as $z \to +1$ holds, and applying this to

$$Q'(z) = mga + \frac{1}{r^2J_1^2J_3}\left\{\frac{S(z)S'(z)}{1-z^2} + \frac{zS(z)^2}{(1-z^2)^2}\right\}.$$

gives

$$U''(0) = -Q'(+1) = -mga + \frac{S'(+1)^2}{4r^2J_1^2J_3}.$$

Another calculation, taking advantage of the condition $S(+1) = 0$ yields

$$S'(+1) = \frac{-C_2rJ_1[J_3 + mr(a+r)]}{J_1 + m(a+r)^2}.$$

If the constant C_2 is not large, U has a local maximum at $\theta = 0$, and for a particular value of the energy, the orbit approaches the vertical position asymptotically. If C_2 is large, there is a local minimum, and the orbit passes periodically through the top position for a range of values for the energy.

Case 4: $S(-1) = 0$ and $S(+1) \neq 0$

The expression

$$\frac{S(z)}{1+z}$$

is analytic at $z = -1$, and U is analytic on the interval $(0, 2\pi)$, with $U(\theta) \to \infty$ as $\theta \to 0, 2\pi$.

Also, $U'(\pi) = 0$ and $U''(\pi) = Q'(-1)$ hold. In this case we have

$$\frac{S(z)}{1-z^2} \to \frac{S'(-1)}{2}$$

as $z \to -1$, which leads to

$$U''(\pi) = Q'(-1) = mga + \frac{S'(-1)^2}{4r^2 J_1^2 J_3}.$$

with

$$S'(-1) = \frac{C_2 r J_1 [J_3 + mr(a-r)]}{J_1 + m(a-r)^2}.$$

Thus U has a local minimum at $\theta = \pi$, and the orbit passes periodically through the bottom position for a range of values for the energy.

Chapter 10

Sphere on a Surface of Revolution

In this section we consider the example of a dynamically symmetric sphere rolling without slipping on a surface of revolution, in the presence of a uniform gravitational field.

10.1 Formulating the Problem

In this subsection we formulate the Lagrangian and the constraint equations for this system, using globally defined quasi-velocities.

Consider figure 10.1 below.

The figure is a cross-section.

We will use the notation of section 2.5 in this section, as in the preceding one. We again take a fixed set of cartesian axes, with associated orthonormal vectors e_1, e_2 and e_3 as before. We take the e_3 axis to be the axis of symmetry.

The distance of the center of mass from the axis of symmetry is denoted by a. The height of the center of mass above the $e_1 \times e_2$ plane is denoted by z.

In figure 10.1 we show the surface upon which the sphere rolls, and also the surface on which the center of mass remains. The cross section of this "center of mass" surface is shown as the dotted curve. We let s denote *arc length* along this curve, so that the curve will be given by

$$\begin{aligned} a &= a(s) \qquad \text{and} \\ z &= z(s). \end{aligned}$$

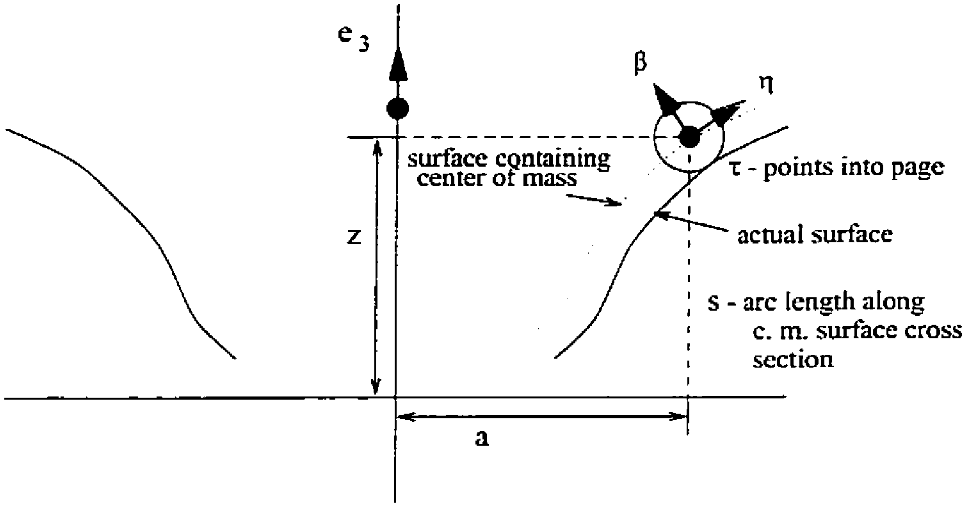


Figure 10.1: Sphere on a Surface of Revolution

in the plane of cross-section. We assume that s may vary over all of $\mathbb{R}$. This may result in a *multiple* covering of the configuration manifold.

If we denote the position of the center of mass by the column vector x, then we may write

$$x(s,\theta) = R_3(\theta)\{a(s)\epsilon_1 + z(s)\epsilon_3\},$$

where θ (not shown in the figure) is the angle the projection of x onto the $\epsilon_1 \times \epsilon_2$ plane makes with ϵ_1. The quantities a, θ and z are cylindrical coordinates for $\mathbb{R}^3$.

In the figure we also show the unit vector

$$\begin{aligned} \eta(s,\theta) &= \frac{\partial x}{\partial s}(s,\theta) \\ &= R_3(\theta)\{a'(s)\epsilon_1 + z'(s)\epsilon_3\} \\ &= R_3(\theta)\eta_0(s), \end{aligned}$$

where we have used

$$\eta_0(s) = \eta(s,0) = a'(s)\epsilon_1 + z'(s)\epsilon_3.$$

We will also use the unit vector

$$\tau(s,\theta) = R_3(\theta)\epsilon_2$$

which is orthogonal to η, and also the unit vector

$$\begin{aligned}
(10.1) \qquad \beta(s,\theta) &= A(\eta(s,\theta))\tau(s,\theta) \\
&= R_3(\theta)A(\eta_0(s))\epsilon_2 \\
&= R_3(\theta)\beta_0(s),
\end{aligned}$$

where we have used

$$\beta_0(s) = \beta(s,0) = A(\eta_0(s))\epsilon_2 = a'(s)\epsilon_3 - z'(s)\epsilon_1.$$

It is important to recognize at this point, that for equation 10.1 to be correct, the choice of parameter s must have a particular orientation. In what follows, the vector β must point to the side of the surface on which the sphere rolls.

We will also find it convenient to introduce the angle ϕ between ϵ_3 and β. This is given by

$$\phi(s) = \arg[a'(s) + iz'(s)].$$

We may calculate

$$\begin{aligned}
\beta^T \frac{\partial \eta}{\partial s} &= [a'(s)\epsilon_3 - z'(s)\epsilon_1]^T[a''(s)\epsilon_1 + z''(s)\epsilon_3] \\
&= a'z'' - z'a'' \\
&= \cos\phi(\sin\phi)' - \sin\phi(\cos\phi)' \\
&= \phi',
\end{aligned}$$

to obtain an expression which we will use later. The (radius of) curvature of the center of mass curve in the figure is given by $1/\phi'$.

As in the preceding section, we also assume there to be a set of orthonormal vectors $\{w_1, w_2, w_3\}$ at the center of mass of the sphere and moving with it, and use the orientation matrix

$$W = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}.$$

We again denote the angular velocity with respect to the fixed axes by ν, so that

$$\dot{W} = A(\nu)W$$

holds. Also, observe that

$$\begin{aligned} \dot{x} &= [R_3(\theta)\{a(s)\epsilon_1 + z(s)\epsilon_3\}]' \\ &= \dot{s}\eta + \dot{\theta}R_3(\theta)A(\epsilon_3)\{a\epsilon_1 + z\epsilon_3\} \\ &= \dot{s}\eta + \dot{\theta}R_3(\theta)\{a\epsilon_2\} \\ &= \dot{s}\eta + a\dot{\theta}\tau \end{aligned}$$

holds.

Let the mass of the sphere be m. We use r to denote the radius of the sphere. Denote the inertia of the body by J. Observe that for a dynamically symmetric sphere, J is just a number. The total kinetic energy of the body is given by

$$\begin{aligned} K &= \frac{1}{2}J\omega^T\omega + \frac{1}{2}m\dot{x}^T\dot{x} \\ &= \frac{1}{2}J\nu^T\nu + \frac{1}{2}m(\dot{s}^2 + a^2\dot{\theta}^2). \end{aligned}$$

The potential energy of the body is given by

$$U = mg\epsilon_3^T x = mgz,$$

where g is the gravitational constant. The Lagrangian for this system is therefore given by

$$L = \frac{1}{2}J\nu^T\nu + \frac{1}{2}m(\dot{s}^2 + a^2\dot{\theta}^2) - mgz. \tag{10.2}$$

The *rolling condition* is that the velocity of the point on the sphere that is in contact with the surface has zero velocity. The constraint therefore is

$$\begin{aligned} 0 &= \dot{x} + A(\nu)(-r\beta) \\ &= \dot{x} + rA(\beta)\nu. \end{aligned} \tag{10.3}$$

In terms of s and θ this becomes

$$\begin{aligned} 0 &= \dot{s} + r\eta^T A(\beta)\nu \\ &= \dot{s} - r\tau^T\nu \end{aligned} \tag{10.4}$$

and

$$\begin{aligned} 0 &= a\dot{\theta} + r\tau^T A(\beta)\nu \\ &= a\dot{\theta} + r\eta^T\nu. \end{aligned} \tag{10.5}$$

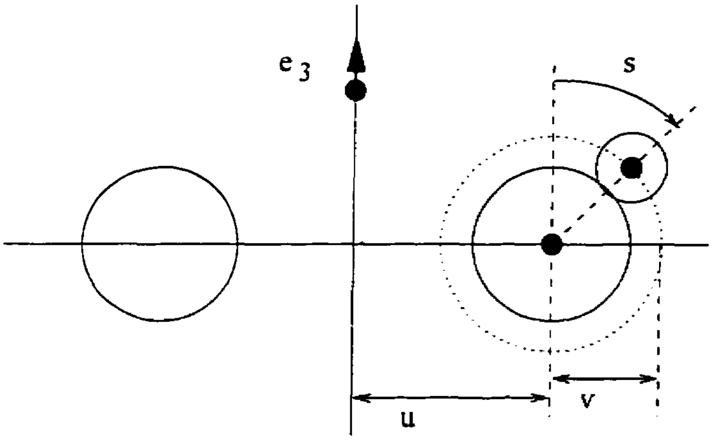


Figure 10.2: Sphere on Horizontal Hoop

As already mentioned, the parameter s must be chosen with a particular orientation. If the orientation is reversed, the vector β will point in the opposite direction. In the constraint above, the surface upon which the sphere rolls is *on the opposite side of the sphere*, although the surface upon which the center of mass remains is unchanged.

The configuration manifold Q, parameterized by (W, s, θ), is 5-dimensional, and the constraint distribution D within TQ is 3-dimensional.

Example. Inside a Vertical Cylinder

For a sphere rolling on the inside of a vertical cylinder, we have a constant, and $z(s) = s$. □

Example. Outside a Horizontal Hoop

For a sphere rolling on (the outside of) a horizontal hoop (see figure 10.2) we have

$$\begin{aligned} a(s) &= u + v\sin\left(\frac{s}{v}\right) \qquad \text{and} \\ z(s) &= v\cos\left(\frac{s}{v}\right). \end{aligned}$$

□

Example. Outside a Sphere

For $u = 0$, the expressions above for a horizontal hoop become

$$\begin{aligned} a(s) &= v \sin\left(\frac{s}{v}\right) \qquad \text{and} \\ z(s) &= v \cos\left(\frac{s}{v}\right). \end{aligned}$$

This is a parameterisation for a sphere, which may be regarded as a surface of revolution. □

10.2 The Group Symmetry

In this section we describe a group symmetry of this system associated with the Lie group

$$G = SO(2) \times SO(3).$$

We will again think of $SO(2)$ as the Lie subgroup $\{R_3(v) | v \in \Re\}$ of $SO(3)$.

To describe the group action, we parameterize G by $\Re \times SO(3)$ using the map

$$(v, M) \mapsto (R_3(v), M).$$

The associated group action on Q is given by the map $(W, s, \theta) \mapsto (\overline{W}, \overline{s}, \overline{\theta})$, with

$$\overline{W} = R_3(v) W M^T, \tag{10.6}$$

$$\overline{s} = s \qquad \text{and} \tag{10.7}$$

$$\overline{\theta} = \theta + v. \tag{10.8}$$

As a consequence of this we immediately have

$$\begin{aligned} \overline{x} &= R_3(v) x, \\ \overline{\eta} &= R_3(v) \eta, \\ \overline{\tau} &= R_3(v) \tau \qquad \text{and} \\ \overline{\beta} &= R_3(v) \beta \end{aligned}$$

We must show that the Lagrangian L is preserved by this group action. Once again we have

$$L = \frac{1}{2} J \nu^T \nu + \frac{1}{2} m(\dot{s}^2 + a^2 \dot{\theta}^2) - mgz.$$

Since s is group invariant, so is z. We also have

$$\dot{\overline{\theta}} = \dot{\theta}.$$

From equation 10.6 we have

$$\begin{aligned} A(\overline{\nu})\overline{W} &= \dot{\overline{W}} \\ &= R_3(\psi)\dot{W} M^T \\ &= R_3(\psi) A(\nu) W M^T \\ &= A(R_3(\psi)\nu)\overline{W}. \end{aligned}$$

It follows that

$$\overline{\nu} = R_3(\psi)\nu$$

holds. Thus we obtain $\overline{L} = L$, and so L is preserved by the group action.

We must also show that the constraint,

$$0 = \dot{x} + rA(\beta)\nu,$$

is preserved. For this we have

$$\begin{aligned} \dot{\overline{x}} + rA(\overline{\beta})\overline{\nu} &= R_3(\psi)\dot{x} + rA(R_3(\psi)\beta)R_3(\psi)\nu \\ &= R_3(\psi)\{\dot{x} + rA(\beta)\nu\}. \end{aligned}$$

Since $R_3(\psi)$ is non-singular, we see that the constraint is preserved by the group action.

10.3 A Basis Aligned with the Symmetry

We now apply the theory of chapter 6 to the current group action, given by equations 10.6, 10.7 and 10.8, which we repeat as

$$\begin{aligned} \overline{W} &= R_3(\psi) W M^T, \\ \overline{s} &= s \qquad\qquad \text{and} \\ \overline{\theta} &= \theta + \psi. \end{aligned}$$

To find a basis of vector fields for the distribution V. of vectors tangent to group orbits, we calculate

$$\left.\frac{\partial\{R_3(v)WM^T, s, \theta + v\}}{\partial v}\right|_{(v,M)=(0,I)} = (A(\epsilon_3)W, 0, 1)$$

and

$$\begin{aligned}\left.\frac{\partial\{R_3(v)WR_i(u)^T, s, \theta + v\}}{\partial u}\right|_{(v,u)=(0,0)} &= (-WA(\epsilon_i), 0, 0)\\ &= (A(-W\epsilon_i)W, 0, 0).\end{aligned}$$

We may map these to $(\nu, \dot{s}, \dot{\theta})$-space to obtain instead

$$\begin{aligned}\frac{\partial}{\partial v} &\mapsto (\epsilon_3, 0, 1) \qquad \text{and}\\ \frac{\partial}{\partial R_i} &\mapsto (-W\epsilon_i, 0, 0),\end{aligned}$$

respectively.

These four vector fields are linearly independent *everywhere*.

To find a basis of vector fields for the constraint distribution D, in $(\nu, \dot{s}, \dot{\theta})$-space, we substitute values for ν in the constraints,

$$\begin{aligned}\dot{s} &= r\tau^T\nu \qquad \text{and}\\ a\dot{\theta} &= -r\eta^T\nu.\end{aligned}$$

For $\nu = \beta$ we obtain

$$(\nu, \dot{s}, \dot{\theta}) = (\beta, 0, 0).$$

For $\nu = a\eta$ we obtain

$$(\nu, \dot{s}, \dot{\theta}) = (a\eta, 0, -r).$$

For $\nu = \tau$ we obtain

$$(\nu, \dot{s}, \dot{\theta}) = (\tau, r, 0).$$

Thus $\{\gamma_1, \gamma_2, \kappa\}$ is a basis for D, with

$$\begin{aligned}\gamma_1 &= (\beta, 0, 0),\\ \gamma_2 &= (a\eta, 0, -r) \qquad \text{and}\\ \kappa &= (\tau, r, 0).\end{aligned}$$

It is easily seen that γ_1 and γ_2 lie in V, and so in $S = D \cap V$. At the same time it is clear that $\kappa \notin V$ holds. So we have

$$S = \operatorname{span}\{\gamma_1, \gamma_2\}.$$

The kinetic energy for this system, once again is

$$K = \frac{1}{2} J \nu^T \nu + \frac{1}{2} m(\dot{s}^2 + a^2 \dot{\theta}^2).$$

The associated inner product, in $(\nu, \dot{s}, \dot{\theta})$-space, is given by

$$K\left((\nu, \dot{s}, \dot{\theta}), (\underline{\nu}, \underline{\dot{s}}, \underline{\dot{\theta}})\right) = J \nu^T \underline{\nu} + m(\dot{s}\underline{\dot{s}} + a^2 \dot{\theta}\underline{\dot{\theta}}).$$

Using this we easily obtain

$$\begin{aligned} K(\kappa, \gamma_1) &= 0 \qquad \text{and} \\ K(\kappa, \gamma_2) &= 0, \end{aligned}$$

so that $\{\kappa\}$ is a basis for $H = D \cap S^{\perp}$.

We have not as yet shown that γ_1, γ_2 and κ are group invariant vector fields. We defer this to the next subsection, in which we also show that this basis satisfies the flatness conditions.

10.4 Group Invariance and the Flatness Conditions

Thinking of γ_1, γ_2 and κ as differentiation operators, we have

$$\begin{aligned} \gamma_1(W, s, \theta) &= (A(\beta)W, 0, 0) \\ \gamma_2(W, s, \theta) &= (A(a\eta)W, 0, -r) \qquad \text{and} \\ \kappa(W, s, \theta) &= (A(\tau)W, r, 0). \end{aligned} \tag{10.9}$$

Proposition 10.4.1. *The vector fields γ_1, γ_2 and κ are group invariant.*

Proof. For γ_1 we calculate

$$\begin{aligned} \gamma_1(\overline{W}, \overline{s}, \overline{\theta}) &= \gamma_1(R_3(\psi)WM^T, s, \theta + \psi) \\ &= (R_3(\psi)A(\beta)WM^T, 0, 0) \\ &= (A(R_3(\psi)\beta)R_3(\psi)WM^T, 0, 0) \\ &= (A(\overline{\beta})\overline{W}, 0, 0). \end{aligned}$$

For γ_2 we calculate

$$\begin{aligned}\gamma_2(\overline{W},\overline{s},\overline{\theta}) &= \gamma_2(R_3(\psi)WM^T,s,\theta+\psi)\\ &= (R_3(\psi)A(a\eta)WM^T,0,-r)\\ &= (A(a\overline{\eta})\overline{W},0,-r).\end{aligned}$$

For κ we calculate

$$\begin{aligned}\kappa(\overline{W},\overline{s},\overline{\theta}) &= \kappa(R_3(\psi)WM^T,s,\theta+\psi)\\ &= (R_3(\psi)A(\kappa)WM^T,r,0)\\ &= (A(\overline{\eta})\overline{W},r,0).\end{aligned}$$

These results may now be compared with equations 10.9. $\square$

Next we find the commutators of these vector fields.

Lemma 10.4.1. *The pair-wise brackets of γ_1, γ_2 and κ (in $(\nu,\dot{s},\dot{\theta})$-space) are*

$$[\gamma_1,\gamma_2] = (-(a+rz')\tau,0,0), \tag{10.10}$$
$$[\kappa,\gamma_1] = (-(1+r\rho')\eta,0,0) \quad \textit{and} \tag{10.11}$$
$$[\kappa,\gamma_2] = (ra'\eta+a(1+r\rho')\beta-rR_3(\theta)e_1,0,0). \tag{10.12}$$

Proof. To obtain equation 10.10 we calculate

$$\begin{aligned}[\gamma_1,\gamma_2](W,s,\theta) &= \gamma_1(A(a\eta)W,0,-r)-\gamma_2(A(\beta)W,0,0)\\ &= (A(a\eta)A(\beta)W,0,0)\\ &\qquad -\left(A\left(-r\frac{\partial\beta}{\partial\theta}\right)W+A(\beta)A(a\eta)W,0,0\right)\\ &= \left(A\left(aA(\eta)\beta+r\frac{\partial\beta}{\partial\theta}\right)W,0,0\right).\end{aligned}$$

or in $(\nu,\dot{s},\dot{\theta})$-space,

$$\begin{aligned}[\gamma_1,\gamma_2] &= \left(aA(\eta)\beta+r\frac{\partial\beta}{\partial\theta},0,0\right)\\ &= (-a\tau+rA(e_3)\beta,0,0)\\ &= (-a\tau+r(\tau^TA(e_3)\beta)\tau,0,0)\\ &= (-a\tau+r(-e_3^T\eta)\tau,0,0)\\ &= (-(a+rz')\tau,0,0).\end{aligned}$$

To obtain equation 10.11 we calculate

$$\begin{aligned}
[\kappa,\gamma_1](W,s,\theta) &= \kappa(A(\beta)W,0,0) - \gamma_1(A(\tau)W,r,0)\\
&= \left(A\left(r\frac{\partial\beta}{\partial s}\right)W + A(\beta)A(\tau)W,0,0\right)\\
&\qquad - (A(\tau)A(\beta)W,0,0)\\
&= \left(A\left(r\frac{\partial\beta}{\partial s} + A(\beta)\tau\right)W,0,0\right).
\end{aligned}$$

or in $(\nu,\dot{s},\dot{\theta})$-space,

$$\begin{aligned}
[\kappa,\gamma_1] &= \left(r\frac{\partial\beta}{\partial s} + A(\beta)\tau,0,0\right)\\
&= \left(r\left(\eta^T\frac{\partial\beta}{\partial s}\right)\eta - \eta,0,0\right)\\
&= \left(-\left(1 + r\beta^T\frac{\partial\eta}{\partial s}\right)\eta,0,0\right)\\
&= (-(1 + r\phi')\eta,0,0).
\end{aligned}$$

To obtain equation 10.12 we calculate

$$\begin{aligned}
[\kappa,\gamma_2](W,s,\theta) &= \kappa(A(a\eta)W,0,-r) - \gamma_2(A(\tau)W,r,0)\\
&= \left(A\left(ra'\eta + ar\frac{\partial\eta}{\partial s}\right)W + A(a\eta)A(\tau)W,0,0\right)\\
&\qquad - \left(A\left(-r\frac{\partial\tau}{\partial\theta}\right)W + A(\tau)A(a\eta)W,0,0\right)\\
&= \left(A\left(ra'\eta + ar\frac{\partial\eta}{\partial s} + r\frac{\partial\tau}{\partial\theta} + aA(\eta)\tau\right)W,0,0\right).
\end{aligned}$$

or in $(\nu,\dot{s},\dot{\theta})$-space,

$$\begin{aligned}
[\kappa,\gamma_2] &= \left(ra'\eta + ar\frac{\partial\eta}{\partial s} + r\frac{\partial\tau}{\partial\theta} + aA(\eta)\tau,0,0\right)\\
&= \left(ra'\eta + ar\left(\beta^T\frac{\partial\eta}{\partial s}\right)\beta + rA(\epsilon_3)\tau + a\beta,0,0\right)\\
&= (ra'\eta + a(1 + r\phi')\beta - rR_3(\theta)\epsilon_1,0,0).
\end{aligned}$$

□

We may interpret these results geometrically. The bracket of two vector fields is a measure of the commutativity of the flows of these vector fields.

For γ_1, the sphere is rotated about the axis perpendicular to the surface. The center of mass and point of contact remain fixed.

For γ_2, the sphere is rolled *out of the page*, with respect to figure 10.1. The center of mass and point of contact move on circles at a fixed height. It is important to note that the rate of rotation varies with s, so that θ varies at the same rate no matter which circle the sphere is on.

For κ, the sphere is rolled *to the right*, with respect to figure 10.1. The center of mass and point of contact stay within the cross-section of the figure.

Lemma 10.4.1 asserts that if two of these virtual motions are applied to the sphere, the change in the *position of the center of mass* will be the same whichever order the motions are applied. It is not hard to see this intuitively.

This however is not so for the change in the orientation of the sphere. It is not easy to visualise intuitively what this change should be. We will depend on lemma 10.4.1 for this information.

Proposition 10.4.2. *The basis $\{\gamma_1, \gamma_2, \kappa\}$ satisfies the flatness conditions, equations 7.2 and 7.3 on page 51, which in this case become*

$$\begin{aligned} K(\gamma_1, [\gamma_1, \gamma_2]) &= 0, \\ K(\gamma_2, [\gamma_1, \gamma_2]) &= 0, \\ K(\kappa, [\kappa, \gamma_1]) &= 0 \qquad \textit{and} \\ K(\kappa, [\kappa, \gamma_2]) &= 0. \end{aligned}$$

Proof. We calculate

$$\begin{aligned} K(\gamma_1, [\gamma_1, \gamma_2]) &= J\beta^T\{-(a + rz')\tau\} \\ &= 0, \\ K(\gamma_2, [\gamma_1, \gamma_2]) &= Ja\eta^T\{-(a + rz')\tau\} \\ &= 0, \\ K(\kappa, [\kappa, \gamma_1]) &= J\tau^T\{-(1 + r\sigma')\eta\} \\ &= 0 \\ &\text{and} \\ K(\kappa, [\kappa, \gamma_2]) &= J\tau^T\{ra'\eta + a(1 + r\sigma')\beta - rR_3(\theta)\epsilon_1\} \\ &= 0. \end{aligned}$$

□

By corollary 8.4.1 on page 68, we conclude that the reduced motion on Q/G is Hamiltonian.

10.5 The Adjoint Equation

In this subsection, we will find the form of the adjoint equation, equation 7.11 on page 57, or

$$\kappa(K(\gamma_i,\gamma_k)\Omega^k) = \{K([\kappa,\gamma_i],\gamma_k) + K(\kappa,[\gamma_k,\gamma_i])\}\Omega^k,$$

for this problem.

We will however use suffixes for the components of Ω, in order to be consistent with the notation used in this section.

We will again work from

$$\begin{aligned}
\gamma_1(W,s,\theta) &= (A(\beta)W,0,0) \\
\gamma_2(W,s,\theta) &= (A(a\eta)W,0,-r) \qquad \text{and} \\
\kappa(W,s,\theta) &= (A(\tau)W,r,0),
\end{aligned}$$

together with

$$\begin{aligned}
[\gamma_1,\gamma_2] &= (-(a+rz')\tau,0,0), \\
[\kappa,\gamma_1] &= (-(1+r\sigma')\eta,0,0) \qquad \text{and} \\
[\kappa,\gamma_2] &= (ra'\eta + a(1+r\sigma')\beta - rR_3(\theta)\epsilon_1,0,0).
\end{aligned}$$

Proposition 10.5.1. *The adjoint equation for the sphere rolling on a surface of revolution becomes*

$$\begin{aligned}
\kappa\Omega_1 &= r(-a\sigma' + z')\Omega_2 \\
&\textit{and} \\
\kappa\{a^2\Omega_2\} &= \frac{Jra\sigma'}{J+mr^2}\Omega_1.
\end{aligned}$$

Proof. We calculate

$$
\begin{aligned}
K(\gamma_1,\gamma_1) &= J\beta^T\beta \\
&= J. \\
K(\gamma_1,\gamma_2) &= J\beta^T(a\eta) \\
&= 0. \\
K(\gamma_2,\gamma_2) &= J(a\eta)^T(a\eta) + ma^2(-r)^2 \\
&= a^2(J + mr^2) \\
K([\kappa,\gamma_1],\gamma_1) &= J\{-(1 + r\sigma')\eta\}^T\beta \\
&= 0. \\
K([\kappa,\gamma_1],\gamma_2) &= J\{-(1 + r\sigma')\eta\}^T(a\eta) \\
&= -aJ(1 + r\sigma'). \\
K([\kappa,\gamma_2],\gamma_1) &= J\{ra'\eta + a(1 + r\sigma')\beta - rR_3(\theta)\epsilon_1\}^T\beta \\
&= J\{a(1 + r\sigma') - r\epsilon_1^T\beta_0\} \\
&= J\{a(1 + r\sigma') + rz'\}. \\
K([\kappa,\gamma_2],\gamma_2) &= J\{ra'\eta + a(1 + r\sigma')\beta - rR_3(\theta)\epsilon_1\}^T(a\eta) \\
&= aJ\{ra' - r\epsilon_1^T\eta_0\} \\
&= 0 \\
&\text{and} \\
K(\kappa,[\gamma_1,\gamma_2]) &= J\tau^T\{-(a + rz')\tau\} \\
&= -J(a + rz').
\end{aligned}
$$

Substituting into the adjoint equation gives

$$
\begin{aligned}
\kappa\{J\Omega_1\} &= J\{-a(1 + r\sigma') + (a + rz')\}\Omega_2 \\
&= Jr(-a\sigma' + z')\Omega_2 \\
&\text{and} \\
\kappa\{a^2(J + mr^2)\Omega_2\} &= J\{[a(1 + r\sigma') + rz'] - (a + rz')\}\Omega_1 \\
&= Jra\sigma'\Omega_1.
\end{aligned}
$$

□

We now observe that since s is group invariant, it induces a coordinate on Q/G. Recalling that $\kappa s = r$ holds, so that κ pushes down to $r\,\partial/\partial s$ on Q/G, we have

Corollary 10.5.1. *Using s as a coordinate for Q/G, the adjoint equation pushes down to*

$$\frac{d\Omega_1}{ds} = \left(-a\frac{d\phi}{ds} + \frac{dz}{ds}\right)\Omega_2 \tag{10.13}$$

and

$$\frac{d}{ds}\{a^2\Omega_2\} = \frac{Ja}{J+mr^2}\frac{d\phi}{ds}\Omega_1. \tag{10.14}$$

on Q/G.

In the case where ϕ' is never zero, we may use ϕ as the independent variable, in place of s. If we introduce the quantity

$$\lambda = \frac{1}{\phi'},$$

which is the (finite radius of) curvature of the center of mass curve, and also replace Ω_2 with $\overline{\Omega}_2 = ra\Omega_2$, we easily obtain:

Corollary 10.5.2. *In terms of Ω_1, $\overline{\Omega}_2$ and ϕ the adjoint equation becomes*

$$\frac{d\Omega_1}{d\phi} = -\frac{\overline{\Omega}_2}{r}\left(1 - \frac{\lambda\sin\phi}{a}\right) \tag{10.15}$$

and

$$\frac{d\overline{\Omega}_2}{d\phi} + \frac{\lambda\cos\phi}{a}\overline{\Omega}_2 = \frac{J}{J+mr^2}r\Omega_1. \tag{10.16}$$

Equations 10.15 and 10.16 are found in Routh [31, Article 230], using force arguments. The variables used there differ slightly from those used here.

Example. Inside a Vertical Cylinder

For a sphere inside a vertical cylinder we have a constant and $z(s) = s$, and so have

$$\phi' = a'z'' - z'a'' = 0.$$

Equations 10.13 and 10.14 become

$$\frac{d\Omega_1}{ds} = \Omega_2 \qquad \text{and}$$
$$\frac{d\Omega_2}{ds} = 0.$$

and are solved by

$$\begin{aligned} \Omega_1 &= C_1 s + C_2 \qquad \text{and} \\ \Omega_2 &= C_1. \end{aligned}$$

□

Example. Outside a Horizontal Hoop

For a sphere rolling on (the outside of) a horizontal hoop (see figure 10.2) we have

$$\begin{aligned} a(s) &= u + v\sin\left(\frac{s}{v}\right) \qquad \text{and} \\ z(s) &= v\cos\left(\frac{s}{v}\right). \end{aligned}$$

and so have

$$\begin{aligned} a' &= \cos\left(\frac{s}{v}\right). \\ z' &= -\sin\left(\frac{s}{v}\right). \\ \sigma' &= a'z'' - z'a'' \\ &= -\frac{1}{v}. \\ a\sigma' &= -\frac{u}{v} - \sin\left(\frac{s}{v}\right) \qquad \text{and} \\ -a\sigma' + z' &= \frac{u}{v}. \end{aligned}$$

Equations 10.13 and 10.14 become

$$\begin{aligned} \frac{d\Omega_1}{ds} &= \frac{u}{v}\Omega_2 \qquad \text{and} \\ \frac{d}{ds}\left\{\left[u + v\sin\left(\frac{s}{v}\right)\right]^2 \Omega_2\right\} &= \frac{J}{J + mr^2}\left\{-\frac{u}{v} - \sin\left(\frac{s}{v}\right)\right\}\Omega_1. \end{aligned}$$

We will examine this example more closely in a later chapter. □

Example. Outside a Sphere

As previously observed. for $u = 0$. the equations for a horizontal hoop become those for a sphere. and the equations above become

$$
\begin{aligned}
\frac{d\Omega_1}{ds} &= 0 \qquad \text{and} \\
\frac{d}{ds}\left\{\left[v \sin\left(\frac{s}{v}\right)\right]^2 \Omega_2\right\} &= -\frac{J}{J + mr^2} \sin\left(\frac{s}{v}\right) \Omega_1.
\end{aligned}
$$

These are solved by

$$
\begin{aligned}
\Omega_1 &= C_1 \qquad \text{and} \\
\Omega_2 &= \left\{\frac{J}{J + mr^2}\frac{C_1}{v} \cos\left(\frac{s}{v}\right) + C_2\right\} \Big/ \left\{\sin^2\left(\frac{s}{r}\right)\right\}.
\end{aligned}
$$

It is interesting to observe that if the ball is rolling inside the sphere. we obtain the same adjoint equation. We place the moving sphere on the inside by replacing the parameter s by $-s$ (see the comments following equation 10.5). But the equations thus obtained are identical. This is not so for the hoop. □

10.6 The Reduced Kinetic Energy and Potential Function

In this subsection. we will find the form of the kinetic energy and the potential function for the reduced problem.

Recall (lemma 8.4.1) that the reduced potential function $\overline{U}$ is determined by

$$\overline{U} \circ \Lambda = U + \frac{1}{2}K(\Omega.\Omega).$$

while the reduced kinetic energy inner product $\overline{K}$ is induced by the requirement (lemma 8.4.2)

$$\overline{K}(T\Lambda(\tau).T\Lambda(\tau)) = K(\tau.\tau).\ \forall \tau \in H \div R.$$

For the sphere on a surface of revolution. the potential function may be written as

$$U(s) = mgz(s).$$

The kinetic energy inner product once again is given, in $(\nu, \dot{s}, \dot{\theta})$-space by

$$K\left((\nu, \dot{s}, \dot{\theta}), (\underline{\nu}, \underline{\dot{s}}, \underline{\dot{\theta}})\right) = J\nu^T\underline{\nu} + m(\dot{s}\underline{\dot{s}} + a^2\dot{\theta}\underline{\dot{\theta}}).$$

We also have

$$\Omega = \Omega_1\gamma_1 + \Omega_2\gamma_2,$$

and the elements of our aligned basis are

$$\begin{aligned} \gamma_1 &= (\beta, 0, 0), \\ \gamma_2 &= (a\eta, 0, -r) \qquad\qquad \text{and} \\ \kappa &= (\tau, r, 0). \end{aligned}$$

First we will examine the kinetic energy. Since

$$\kappa s = r$$

holds, when κ is thought of as a differentiation operator, κ is the horizontal lift of $r\partial/\partial s$. Therefore we have

$$\begin{aligned} \overline{K}\left(\frac{\partial}{\partial s}, \frac{\partial}{\partial s}\right) &= \frac{1}{r^2}K(\kappa, \kappa) \\ &= \frac{J + mr^2}{r^2}. \end{aligned}$$

Expressed as a function of s then, the reduced kinetic energy is

$$\overline{K}(s, \dot{s}) = \frac{1}{2}\left(m + \frac{J}{r^2}\right)\dot{s}^2.$$

Next we examine the potential energy. We begin with

$$\begin{aligned} \overline{U}(s) &= mgz(s) + \frac{1}{2}\{\Omega_1(s)^2 K(\gamma_1, \gamma_1) \\ &\qquad + 2\Omega_1(s)\Omega_2(s)K(\gamma_1, \gamma_2) + \Omega_2(s)^2 K(\gamma_2, \gamma_2)\} \\ &= mgz(s) + \frac{1}{2}\{J\Omega_1(s)^2 + (J + mr^2)a(s)^2\Omega_2(s)^2\}. \end{aligned}$$

To proceed further, we need an actual solution to the adjoint equation. So we do this for examples.

Example. Inside a Vertical Cylinder

For a sphere inside a vertical cylinder we have a constant and $z(s) = s$, and our solution to the adjoint equation was

$$\begin{aligned} \Omega_1 &= C_1 s + C_2 \qquad \text{and} \\ \Omega_2 &= C_1. \end{aligned}$$

So we obtain

$$\overline{U}(s) = mgs + \frac{1}{2}\{J[C_1 s + C_2]^2 + (J + mr^2)a^2 C_1^2\}.$$

The reduced motion is easily understood. If $C_1 = 0$ holds, then we have

$$\overline{U}(s) = mgs + C_3.$$

Otherwise we have

$$\overline{U}(s) = \frac{1}{2} J C_1^2 (s - s_0)^2 + C_3.$$

In each case, C_3 is a constant, determined by the initial value of $\dot{s}$, and s_0 satisfies

$$s_0 = -\frac{mg + JC_1 C_2}{JC_1^2}$$

in the second case.

In the first case the ball simply rolls straight down the cylinder. In the second, it rises and falls *harmonically* about the height given by s_0. □

Example. Outside a Sphere

For a sphere we have

$$\begin{aligned} a(s) &= v \sin\left(\frac{s}{v}\right) \qquad \text{and} \\ z(s) &= v \cos\left(\frac{s}{v}\right). \end{aligned}$$

and our solution to the adjoint equation was

$$\begin{aligned} \Omega_1 &= C_1 \qquad \text{and} \\ \Omega_2 &= \left\{\frac{J}{J + mr^2}\frac{C_1}{v}\cos\left(\frac{s}{v}\right) + C_2\right\} \Big/ \left\{\sin^2\left(\frac{s}{v}\right)\right\}. \end{aligned}$$

So we obtain

$$\begin{aligned}\overline{U}(s) &= mgv\cos\left(\frac{s}{v}\right) + \frac{1}{2}JC_1^2 \\ &\quad + \frac{1}{2}(J+mr^2)v^2 \\ &\quad \left\{\frac{JC_1}{(J+mr^2)v}\cos\left(\frac{s}{v}\right) + C_2\right\}^2 \Big/ \left\{\sin\left(\frac{s}{v}\right)\right\}^2 .\end{aligned}$$

As for the adjoint equation, this result is unchanged if s is replaced by $-s$. So we obtain the same result if the moving sphere is rolling inside the fixed sphere.

If we use ϕ as the independent variable in place of s, for the total reduced energy we have

$$\begin{aligned}E &= \frac{1}{2}\left(m + \frac{J}{r^2}\right)v^2\dot{\phi}^2 + mgv\cos\phi + \frac{1}{2}JC_1^2 \\ &\quad + \frac{1}{2}(J+mr^2)v^2\frac{1}{\sin^2\phi}\left\{\frac{JC_1}{(J+mr^2)v}\cos\phi + C_2\right\} .\end{aligned}$$

This is identical in form to that found for the **Lagrange top**. In fact, it is not hard to establish that the possible trajectories for the center of mass of the rolling sphere on the fixed sphere, are precisely the same as those for the tip of a Lagrange top, which is of course also constrained to lie on a sphere. □

Chapter 11

Discussion of Basketball Rolling on Its Hoop

In this section, we will give an examination of the reduced system for the ball on a hoop, for the special case of a basketball rolling on a basketball hoop.

Recall figure 10.2 on page 113. There, our parameterization was

$$\begin{aligned} a(s) &= u + v \sin\left(\frac{s}{v}\right) \qquad \text{and} \\ z(s) &= v \cos\left(\frac{s}{v}\right). \end{aligned}$$

We showed that the kinetic energy of the reduced system is given by

$$K(s, \dot{s}) = \frac{1}{2}\left(m + \frac{J}{r^2}\right)\dot{s}^2,$$

and the reduced potential energy by

$$U(s) = mgz(s) + \frac{1}{2}\{J\Omega_1(s)^2 + (J + mr^2)a(s)^2\Omega_2(s)^2\},$$

where Ω_1 and Ω_2 are solutions to the adjoint equation,

$$\begin{aligned} \frac{d\Omega_1}{ds} &= \frac{u}{v}\Omega_2 \qquad \text{and} \\ \frac{d}{ds}\{a(s)^2\Omega_2\} &= \frac{-J}{v(J + mr^2)}a(s)\Omega_1. \end{aligned}$$

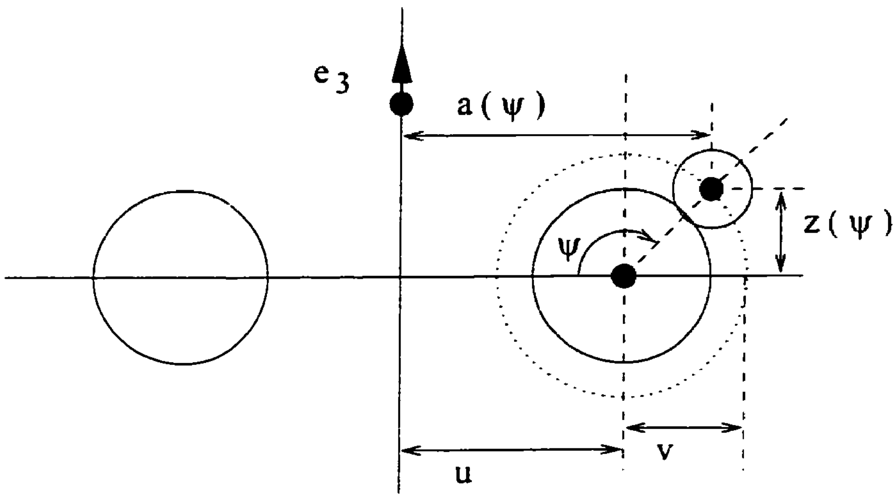


Figure 11.1: Sphere on Horizontal Hoop

New Parameterization

For physical reasons, which will become clear, we make the change of variable

$$s = v\left(\psi - \frac{\pi}{2}\right).$$

We refer the reader to figure 11.1.

The equations above become

$$\begin{aligned} a(\psi) &= u - v\cos\psi, \\ z(\psi) &= v\sin\psi, \\ K(\psi,\dot{\psi}) &= \frac{(J+mr^2)v^2}{2r^2}\dot{\psi}^2, \qquad \text{and} \\ U(\psi) &= mgz(\psi) + \frac{1}{2}\{J\Omega_1(\psi)^2 + (J+mr^2)a(\psi)^2\Omega_2(\psi)^2\}. \end{aligned}$$

with Ω_1 and Ω_2 solutions to

(11.1)
$$\begin{aligned} \frac{d\Omega_1}{d\psi} &= u\Omega_2 \qquad \text{and} \\ \frac{d}{d\psi}\{a(\psi)^2\Omega_2\} &= \frac{-J}{J+mr^2}a(\psi)\Omega_1. \end{aligned}$$

The shape of the function U will determine the possible reduced motions. We are particularly interested in trajectories passing through $w = 0$.

At the Horizontal Position

To gain some insight we will find the values of $U'(0)$ and $U''(0)$, in terms of $C_1 = \Omega_1(0)$ and $C_2 = \Omega_2(0)$. We begin by finding

$$\begin{aligned} a'(w) &= v \sin w, \\ z'(w) &= v \cos w, \\ a''(w) &= v \cos w \qquad \text{and} \\ z''(w) &= -v \sin w, \end{aligned}$$

which give

$$\begin{aligned} a(0) &= u - v, \\ z(0) &= 0, \\ a'(0) &= 0, \\ z'(0) &= v, \\ a''(0) &= v \qquad \text{and} \\ z''(0) &= 0. \end{aligned}$$

From the adjoint equation we have

$$\begin{aligned} \Omega_1' &= u\Omega_2, \\ a\Omega_2' &= -2a'\Omega_2 - \frac{J}{J + mr^2}\Omega_1, \\ \Omega_1'' &= u\Omega_2' \qquad \text{and} \\ a\Omega_2'' &= -3a'\Omega_2' - 2a''\Omega_2 - \frac{J}{J + mr^2}\Omega_1', \end{aligned}$$

giving

$$\begin{aligned} \Omega_1'(0) &= uC_2, \\ \Omega_2'(0) &= -\frac{J}{(u - v)(J + mr^2)}C_1, \\ \Omega_1''(0) &= -\frac{uJ}{(u - v)(J + mr^2)}C_1 \qquad \text{and} \\ \Omega_2''(0) &= -\frac{uJ + 2v(J + mr^2)}{(u - v)(J + mr^2)}C_2. \end{aligned}$$

Finally then, differentiating the potential, we have

$$U' = mgz' + J\Omega_1\Omega_1' + (J + mr^2)\{aa'\Omega_2^2 + a^2\Omega_2\Omega_2'\}$$

and

$$U'' = mgz'' + J\{\Omega_1\Omega_1'' + (\Omega_1')^2\} + (J + mr^2)$$
$$\{aa''\Omega_2^2 + (a')^2\Omega_2^2 + 4aa'\Omega_2\Omega_2' + a^2(\Omega_2')^2 + a^2\Omega_2\Omega_2''\}.$$

giving

$$U'(0) = v(mg + JC_1C_2) \qquad \text{and}$$

$$\begin{aligned}
U''(0) &= J\left\{-\frac{uJC_1^2}{(u-v)(J+mr^2)} + u^2C_2^2\right\} + (J + mr^2) \\
&\quad \left\{(u-v)vC_2^2 + (u-v)^2\left[-\frac{J}{(u-v)(J+mr^2)}C_1\right]^2\right. \\
&\quad \left. +(u-v)^2C_2\left[-\frac{uJ + 2v(J+mr^2)}{(u-v)(J+mr^2)}C_2\right]\right\} \\
&= v\left\{-C_1^2\frac{J^2}{(u-v)(J+mr^2)} + C_2^2[vJ - mr^2(u-v)]\right\}.
\end{aligned}$$

It is clear that C_1 and C_2 may be chosen so as to achieve any value for the quantity $\gamma = U'(0)$. For the case $U'(0) < 0$, for a range of values of the energy, the trajectory is such that the center of the ball drops below the horizontal plane of the hoop, and rises again.

Now $0 < v < u$ must hold. We see that any value for $U''(0)$ may not be positive, provided that the coefficient of C_2^2 above is negative, or

$$vJ - mr^2(u - v) < 0 \Rightarrow v < \frac{mr^2u}{J + mr^2}.$$

A Basketball

For a basketball, the mass may be regarded as distributed evenly on its surface. A straight forward calculation yields

$$J = \frac{2mr^2}{3},$$

and the requirement above becomes

$$v < \frac{3u}{5}.$$

For a basketball, the circumference of the ball is about 30 inches, the inner diameter of the hoop (or diameter of the hole through which the ball passes) is about 18 inches, and the thickness of the hoop is about 3/4 of an inch. Using centimeters as the unit of length, we have $r \approx 12.1276$, $a \approx 23.8125$ and $c \approx 13.0801$. We thus have $3a/5 \approx 14.2875$, and so $U''(0)$ may not be positive.

If we have $U'(0) = 0$ with $U''(0) < 0$, then there is a local maximum of the potential at $w = 0$. If $\Omega_1(0)$ and $\Omega_2(0)$ are varied from these values such that $U'(0)$ becomes negative, the local maximum moves lower.

Numerical Investigation

We will now describe, using a number of plots produced by the Maple mathematical software, a numerical investigation into this effect.

In all of the plots discussed below, the units of w have been adjusted to units of π radians.

We first require a solution to equations 11.1, in terms of C_1 and C_2. For the $C_1 = 1$ and $C_2 = 0$ solution we will write $\Omega_1 = f_{11}(w)$ and $\Omega_2 = f_{12}(w)$. For the $C_1 = 0$ and $C_2 = 1$ solution we will write $\Omega_1 = f_{21}(w)$ and $\Omega_2 = f_{22}(w)$.

It is easily seen from equations 11.1 that f_{11} and f_{22} are even functions, while f_{12} and f_{21} are odd functions.

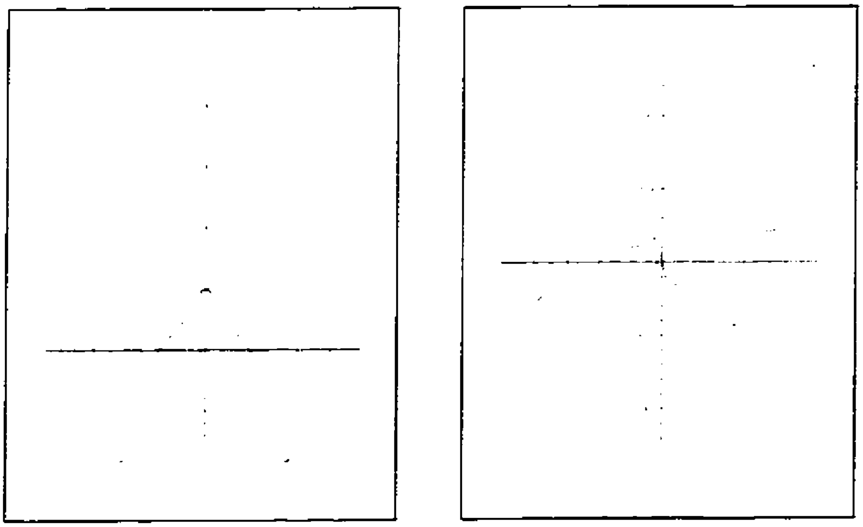
Figure 11.2 shows a plot of the numerical solution obtained for $C_1 = 1$ and $C_2 = 0$ (that is, of f_{11} and f_{12}). The plot has been done for the range -4 to 4 (that is, -4π to $+4\pi$). The tickmarks on the independent axis are at -4, -2, 0, $+2$ and $+4$.

Figure 11.3 shows a plot of the numerical solution obtained for $C_1 = 0$ and $C_2 = 1$ (that is, of f_{21} and f_{22}). Again, the tickmarks on the independent axis are at -4, -2, 0, $+2$ and $+4$.

It is apparent from these plots that the solution to the adjoint equation is not periodic.

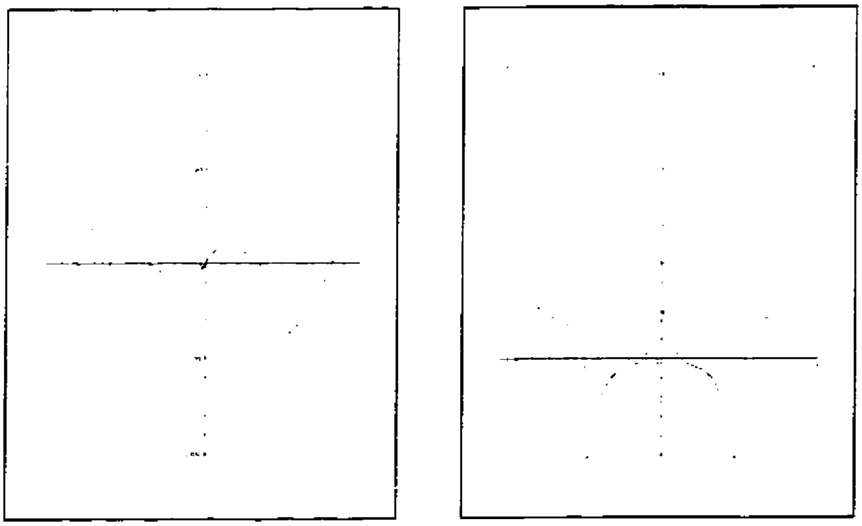
Using these fundamental solutions, we have

$$
\begin{aligned}
U(w) &= mgz(w) + \frac{1}{2}\{J[C_1 f_{11}(w) + C_2 f_{21}(w)]^2 \\
&\qquad + (J + mr^2)a(w)^2[C_1 f_{12}(w) + C_2 f_{22}(w)]^2\} \\
&= mgz(w) + \frac{1}{2}m\{A(w)C_1^2 + 2B(w)C_1 C_2 + C(w)C_2^2\}.
\end{aligned}
$$



Plot of $\Omega_1 (f_{11})$ Plot of $\Omega_2 (f_{12})$

Figure 11.2: Plot of Solution for $\Omega_1(0) = 1$ and $\Omega_2(0) = 0$



Plot of $\Omega_1 (f_{21})$ Plot of $\Omega_2 (f_{22})$

Figure 11.3: Plot of Solution for $\Omega_1(0) = 0$ and $\Omega_2(0) = 1$

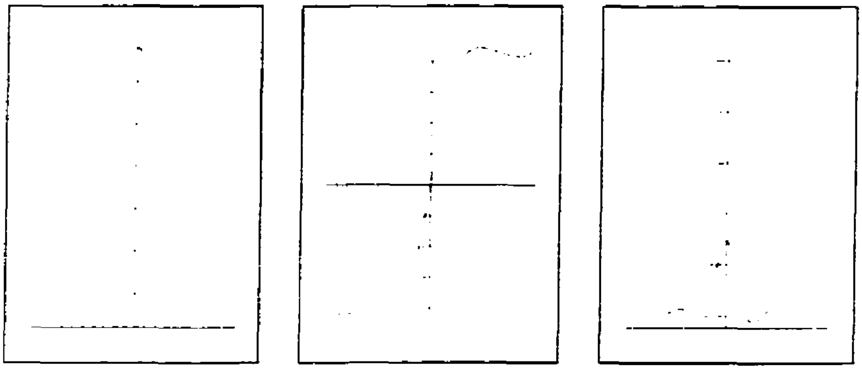


Figure 11.4: Plot of Coefficients A, B and C in Reduced Potential

with

$$\begin{aligned}
A(v) &= \frac{r^2}{3}[2f_{11}(v)^2 + 5a(v)^2 f_{12}(v)^2], \\
B(v) &= \frac{r^2}{3}[2f_{11}(v)f_{21}(v) + 5a(v)^2 f_{12}(v)f_{22}(v)] \quad \text{and} \\
C(v) &= \frac{r^2}{3}[2f_{21}(v)^2 + 5a(v)^2 f_{22}(v)^2].
\end{aligned}$$

The functions A and C are even, while the function B is odd.

Figure 11.4 shows a plot of the numerical solution obtained for A, B and C. The plot has been done for the range -1 to $+1$ (that is, $-\pi$ to $+\pi$). The tickmarks on the independent axis are at -1, -0.5, 0, $+0.5$ and $+1$.

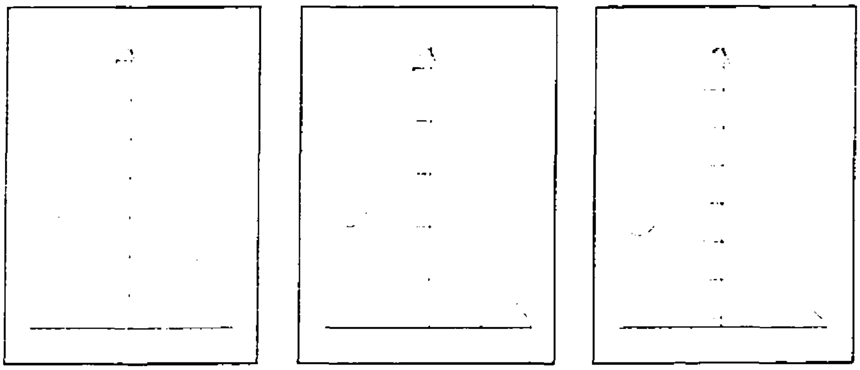
Observe that the mass, m, cancels out of the equations of motion, and so may be set to 1. Using seconds for the unit of time (we have already chosen centimeters for the unit of length), we have $g \approx 980.66$.

We now examine the shape of the reduced potential function U, for reasonable values of the parameters C_1 and C_2. We first need to consider the physical meaning of these parameters.

Recall from section 10.3, that the basis vectors for $S = V \cap D$ were

$$\begin{aligned}
\gamma_1 &= (\beta, 0, 0) \qquad \text{and} \\
\gamma_2 &= (a\eta, 0, -r),
\end{aligned}$$

in $(\nu, \dot{s}, \dot{\theta})$-space. The symbol β appearing in this equation represented the unit vector normal to the surface. The symbol η represented the unit vector



C_1=-40. C_2=.5 C_1=-50. C_2=.5 C_1=-60. C_2=.5

Figure 11.5: Plots of Reduced Potential for $C_2 = 0.5$

pointing along the longitude of the surface (recall figure 10.1 on page 110). This leads to

$$\begin{aligned} \nu_n(\psi) &= \Omega_1(\psi) \quad \text{and} \\ \dot{\theta}(\psi) &= -r\Omega_2(\psi). \end{aligned}$$

where $\nu_n = \beta^T\nu$ is the "normal component" of the angular velocity.

A value of 2π radians/second for $\dot{\theta}$ corresponds to the ball going once around the hoop per second. The corresponding value for Ω_2 is $2\pi/r \approx 0.5181$. So values for C_2 of 0.5 and 1.0 seem physically reasonable.

A value of 2π radians/second for ν_n would correpond to the ball spinning around the normal direction once per second. The corresponding value for Ω_1 is $2\pi \approx 6.283$. So values for C_1 of from 20 to 60 seem physically reasonable.

Figure 11.5 shows plots of U for $C_2 = 0.5$ and $C_1 = -40, -50$ and -60.

Figure 11.6 shows plots of U for $C_2 = 1.0$ and $C_1 = -20, -30$ and -40.

Figure 11.7 shows a larger version of the plot of U for $C_2 = 1$ and $C_1 = -40$. From this plot, we see that the local maximum has been shifted to $\psi \approx -0.12\pi$, or about 21 degrees below the horizontal position. There are trajectories which cross $\psi = 0$ with $\dot{\psi} < 0$, approach the local maximum, fail to reach it, and then change direction, passing back through the horizontal position. From the shape of the plot of U, we see that these trajectories do not again pass the horizontal position. The basketball has gone below the horizontal, but has popped back out again.

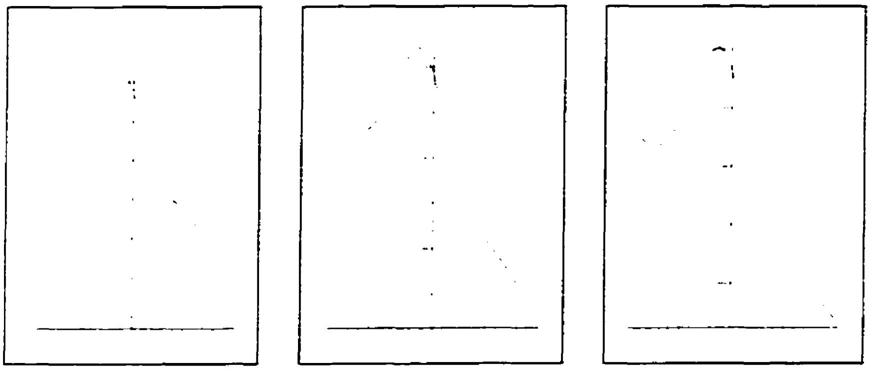


Figure 11.6: Plots of Reduced Potential for $C_2 = 1.0$

There are trajectories of this sort that come arbitarily close to the local maximum position. There is also one which approaches this position asymptotically, taking forever to reach it.

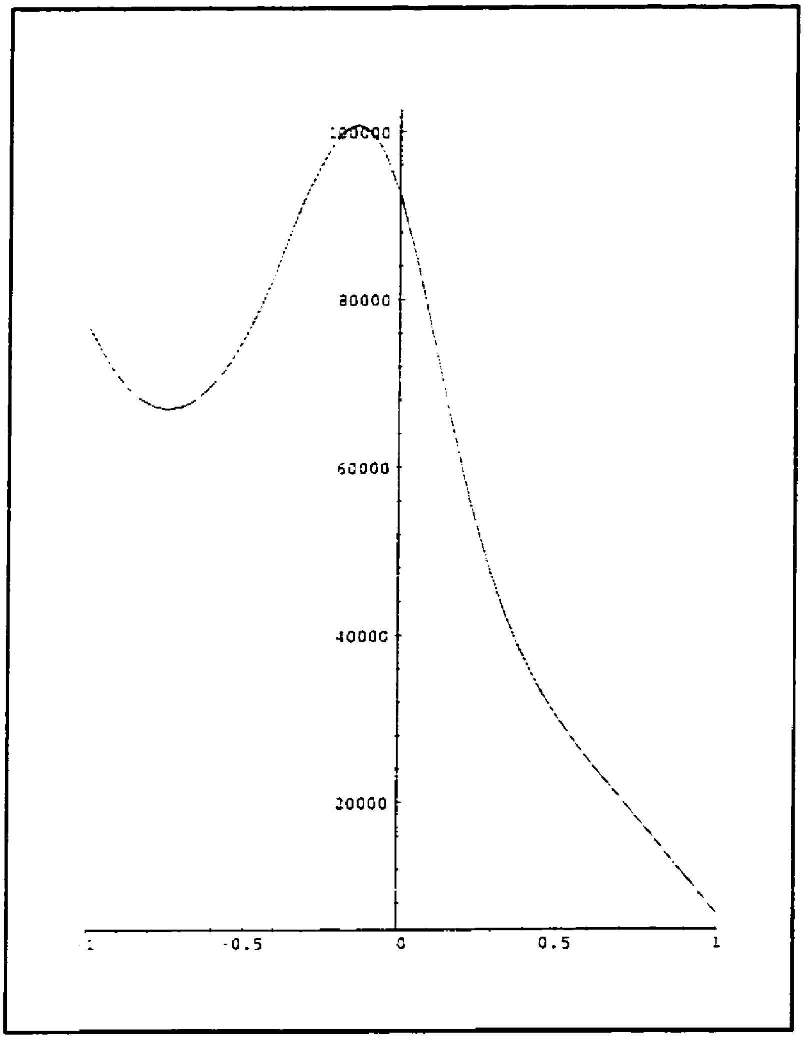
For the trajectories which reverse their direction, we could calculate numerically the number of times that the ball goes around the hoop, while below the horizontal position. The closer the trajectory approaches the local maximum, the larger this number would be. For the asymptotic trajectory this number would be infinite. It is not meaningful, therefore, using this model, to ask the question "How many times does the basketball go around the hoop before popping back out?".

In reality of course, the basketball would not go around the hoop forever. There is dissipative friction acting. The basketball would eventually fall through the hoop. Any orbit that attempts to approach too close to the local maximum will also fall through the hole.

Normal Component of Force

The limits of this model are also exceeded if the normal component of the force between the ball and the hoop becomes non-positive during the motion. In such a case the ball will leave the hoop. We will find an expression for this force.

Refering once again to figure 10.1 on page 110, for this problem, the vectors β (normal to the surface), η (along the longitude) and τ (into the



C 1=-40. C 2 =1

Figure 11.7: Plot of Reduced Potential for $C_1 = -40.0$ and $C_2 = 1.0$

figure) may be written (see also figure 11.1)

$$\begin{aligned} \beta &= -R_3(\theta)R_2(\psi)\epsilon_1, \\ \eta &= R_3(\theta)R_2(\psi)\epsilon_3 \qquad \text{and} \\ \tau &= R_3(\theta)\epsilon_2. \end{aligned}$$

and we have

$$\begin{aligned} \dot\theta(\psi) &= -r\Omega_2(\psi) \qquad \text{and} \\ \dot x &= a(\psi)\dot\theta(\psi)\tau + v\dot\psi\eta. \end{aligned}$$

The normal force is given by

$$F_n = \beta^T\{m\ddot x + mg\epsilon_3\}.$$

We may calculate

$$\beta^T\epsilon_3 = -\epsilon_3^T R_2(\psi)\epsilon_1 = -\epsilon_3^T\{\cos\psi\epsilon_1 - \sin\psi\epsilon_3\} = \sin\psi$$

and

$$\begin{aligned} \beta^T\ddot x &= (\beta^T\dot x)' - (\beta')^T\dot x \\ &= -(\beta')^T\dot x \\ &= -\{-\dot\theta(\psi)R_3(\theta)A(\epsilon_3)R_2(\psi)\epsilon_1 - \dot\psi R_3(\theta)R_2(\psi)A(\epsilon_2)\epsilon_1\}^T \\ &\qquad \{a(\psi)\dot\theta(\psi)R_3(\theta)\epsilon_2 + v\dot\psi R_3(\theta)R_2(\psi)\epsilon_3\} \\ &= \{\dot\theta(\psi)A(\epsilon_3)R_2(\psi)\epsilon_1 - \dot\psi R_2(\psi)\epsilon_3\}^T\{a(\psi)\dot\theta(\psi)\epsilon_2 + v\dot\psi R_2(\psi)\epsilon_3\} \\ &= -a(\psi)\dot\theta(\psi)^2\epsilon_1^T R_2(\psi)^T A(\epsilon_3)\epsilon_2 - v\dot\psi^2 \\ &\qquad - v\dot\theta(\psi)\dot\psi\epsilon_1^T R_2(\psi)^T A(\epsilon_3)R_2(\psi)\epsilon_3 \\ &= a(\psi)\dot\theta(\psi)^2\epsilon_1^T R_2(\psi)\epsilon_1 - v\dot\psi^2 + v\dot\theta(\psi)\dot\psi\epsilon_1^T A(\epsilon_3)R_2(\psi)^T\epsilon_3 \\ &= a(\psi)\dot\theta(\psi)^2\cos\psi - v\dot\psi^2 - v\dot\theta(\psi)\dot\psi\epsilon_2^T R_2(\psi)^T\epsilon_3 \\ &= a(\psi)\dot\theta(\psi)^2\cos\psi - v\dot\psi^2. \end{aligned}$$

The normal force is

$$F_n = m\{a(\psi)\dot\theta(\psi)^2\cos\psi - v\dot\psi^2 + g\sin\psi\}.$$

This is what one would obtain by resolving components of the two centrifugal forces and the gravitational force. In terms of Ω_2 this becomes

$$\begin{aligned} \frac{F_n}{m} &= a(\psi)r^2\Omega_2(\psi)^2\cos\psi - v\dot\psi^2 + g\sin\psi \\ &= a(\psi)r^2[C_1f_{12}(\psi) + C_2f_{22}(\psi)]^2\cos\psi + g\sin\psi - v\dot\psi^2 \\ &= v\{N(\psi) - \dot\psi^2\}. \end{aligned}$$

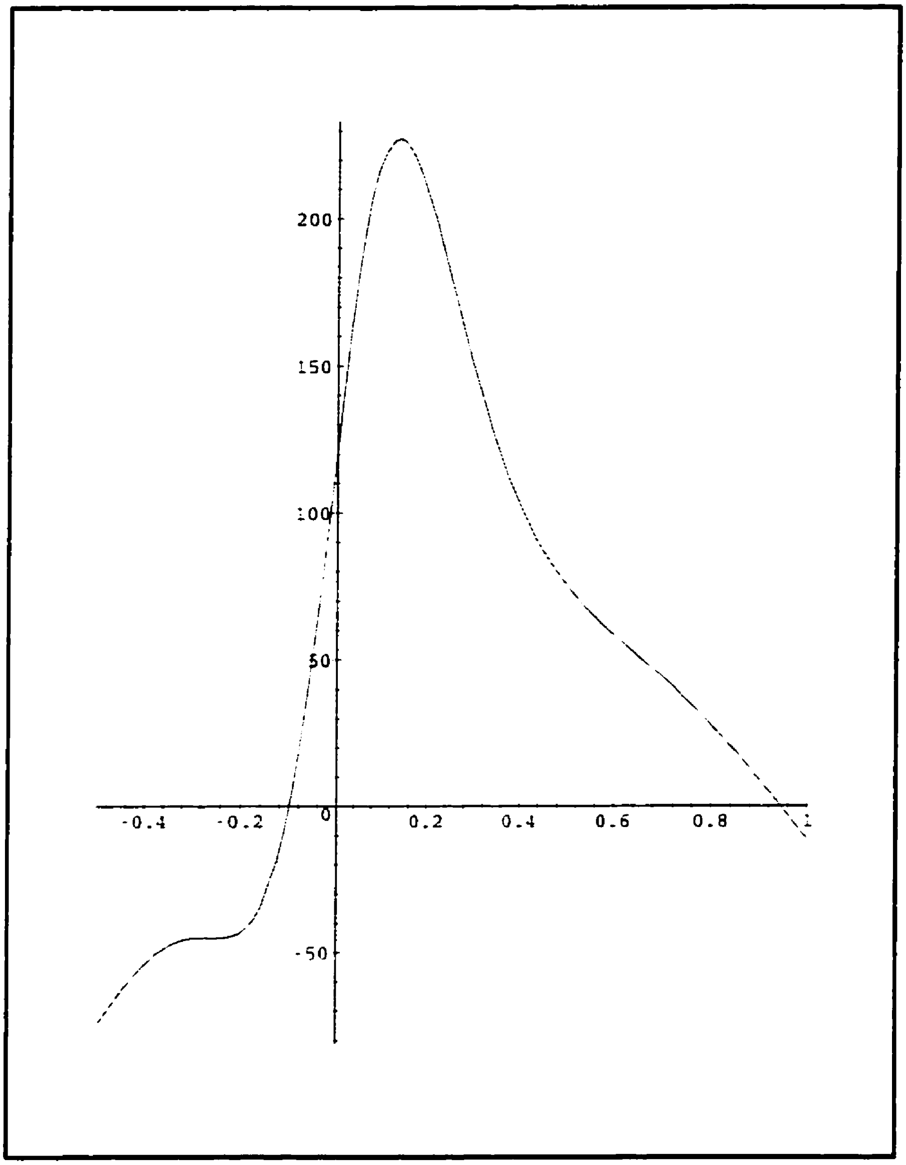
with

$$\mathcal{N}(\psi) = \frac{a(\psi)r^2[C_1 f_{12}(\psi) + C_2 f_{22}(\psi)]^2 \cos\psi + g \sin\psi}{\psi}.$$

The ball remains on the hoop so long as the condition $|\dot{\psi}|^2 < \mathcal{N}(\psi)$ is satisfied. For the trajectories we are interested in, $\dot{\psi}$ is small in magnitude, in fact asymptotically zero for the asymptotic trajectory. Hence we have

$$F_n \approx m\psi\mathcal{N}(\psi).$$

Figure 11.8 shows a plot of $\mathcal{N}$ for $C_2 = 1$ and $C_1 = -40$. From this plot, we see that the ball leaves the hoop for $\psi \approx -0.10\pi$, or about 18 degrees below the horizontal position.



C 1=-40, C 2 =1

Figure 11.8: Plot of Limit for $|\dot{v}|^2$ for $C_1 = -40.0$ and $C_2 = 1.0$

Chapter 12

Three-Wheel Cart

In this section we consider the example of a three wheeled cart rolling on a horizontal plane.

12.1 Formulating the Problem

In this subsection we formulate the Lagrangian and the constraint equations for this system.

Consider figure 12.1 below.

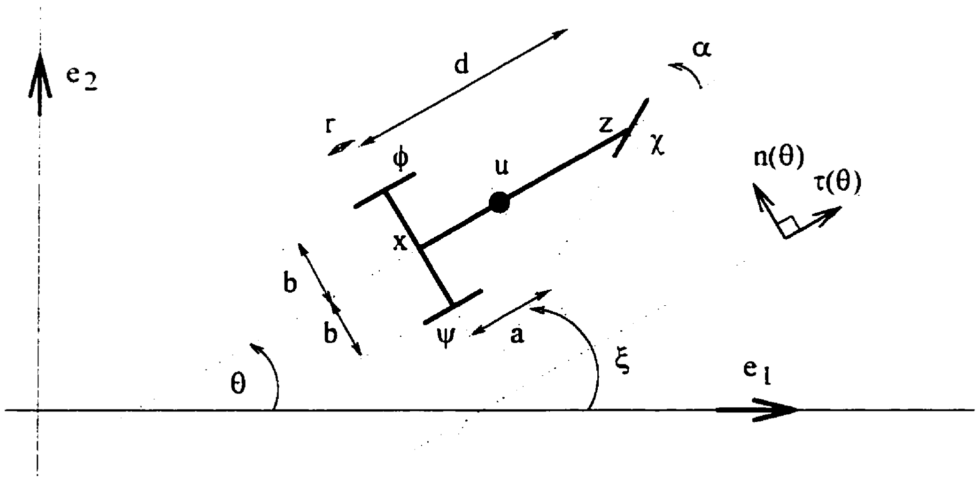


Figure 12.1: Three-Wheeled Cart

The figure is a top view.

There are three identical wheels, except that the front wheel may turn. The radius of each wheel is r. The moment of inertia of a wheel about a perpendicular line through its center is J_W. About a line through a diameter the moment is J_D. The angle that the front wheel makes with the axis of the cart is ϕ, as shown.

The wheels are labeled ϕ, ψ and χ in the figure. These symbols will also represent the angle of rotation of each wheel.

The distances a, b and d are as shown.

The figure shows a set of cartesian axes, which we associate with orthonormal vectors ϵ_1 and ϵ_2 in the usual way.

The symbol x labels the center of the rear axis, and is also the column vector giving its position in the plane. The symbol u is similarly associated with the center of mass of the cart. The symbol z is the column vector giving the position of the front wheel. The moment of inertia of the cart about a vertical axis through u is J. The total mass of the cart is m.

The angle that the axis of the cart makes with the ϵ_1-axis is denoted θ, as shown. The angle that the front wheel makes with the ϵ_1-axis is denoted ξ.

We also introduce the rotation matrix R, defined for a general angle ω by

$$R(\omega) = \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix}.$$

along with the vectors

$$\begin{aligned} \tau(\omega) &= R(\omega)\epsilon_1 \\ &= \begin{bmatrix} \cos\omega \\ \sin\omega \end{bmatrix} \qquad \text{and} \\ n(\omega) &= R(\omega)\epsilon_2 \\ &= \begin{bmatrix} -\sin\omega \\ \cos\omega \end{bmatrix}. \end{aligned}$$

There is no potential energy, so for the kinetic energy K and Lagrangian L we have

$$\begin{aligned} L &= K \\ &= \frac{1}{2}J\dot\theta^2 + \frac{1}{2}m\dot u^T\dot u + \frac{1}{2}J_W(\dot\phi^2 + \dot\psi^2 + \dot\chi^2) + \frac{1}{2}J_D\dot\xi^2. \end{aligned} \tag{12.1}$$

Observe that

$$\begin{aligned} \alpha &= \xi - \theta, \\ z &= x + d\tau(\theta) \qquad\qquad \text{and} \\ u &= x + a\tau(\theta) \end{aligned}$$

hold.

The configuration manifold Q, which may be parameterized by $(\theta, u, \phi, \psi, \chi, \xi)$, is 7-dimensional.

We determime the nonholonomic constraints by considering each wheel. At ϕ we have

$$\begin{aligned} r\dot{\phi}\tau(\theta) &= [x + b\,n(\theta)]' \\ &= \dot{x} - b\dot{\theta}\tau(\theta). \end{aligned}$$

At ψ we obtain

$$r\dot{\psi}\tau(\theta) = \dot{x} + b\dot{\theta}\tau(\theta).$$

At χ we obtain

$$\dot{z} = r\dot{\chi}\tau(\xi),$$

or

$$\begin{aligned} r\dot{\chi}R(\xi)e_1 &= [x + d\tau(\theta)]' \\ &= \dot{x} + d\dot{\theta}\,n(\theta). \end{aligned}$$

From the equation at either ϕ or ψ we obtain

$$n(\theta)^T\dot{x} = 0,$$

and from the equation at χ,

$$\begin{aligned} \tau(\theta)^T\dot{x} &= \tau(\theta)^T[r\dot{\chi}R(\xi)e_1] \\ &= r\dot{\chi}e_1^T R(\theta)^T R(\xi)e_1 \\ &= r\dot{\chi}e_1^T R(\alpha)e_1 \\ &= r\dot{\chi}\cos\alpha. \end{aligned}$$

So we obtain

$$\dot{x} = r\dot{\chi} \cos\alpha \, \tau(\theta).$$

Also from the equation at χ we find

$$\begin{aligned} d\dot{\theta} &= n(\theta)^T [r\dot{\chi} R(\xi) e_1] \\ &= r\dot{\chi} e_2^T R(\theta)^T R(\xi) e_1 \\ &= r\dot{\chi} e_2^T R(\alpha) e_1 \\ &= r\dot{\chi} \sin\alpha. \qquad (12.2) \end{aligned}$$

and from this

$$\begin{aligned} d\dot{u} &= d[x + a\tau(\theta)]' \\ &= r\dot{\chi} d \cos\alpha \, \tau(\theta) + a d\dot{\theta} \, n(\theta) \\ &= r\dot{\chi} \, [d \cos\alpha \, \tau(\theta) + a \sin\alpha \, n(\theta)] \\ &= r\dot{\chi} R(\theta) \lambda(\alpha). \qquad (12.3) \end{aligned}$$

with

$$\lambda(\alpha) = \begin{bmatrix} d\cos\alpha \\ a\sin\alpha \end{bmatrix}.$$

Finally, from the equation at ϕ, we have

$$\begin{aligned} dr\dot{\phi} &= d[\tau(\theta)^T \dot{x} - b\dot{\theta}] \\ &= r\dot{\chi} [d\cos\alpha - b\sin\alpha] \end{aligned}$$

or

$$d\dot{\phi} = \dot{\chi} [d\cos\alpha - b\sin\alpha]. \qquad (12.4)$$

For ψ this becomes

$$d\dot{\psi} = \dot{\chi} [d\cos\alpha + b\sin\alpha]. \qquad (12.5)$$

The constraint equations are equations 12.2, 12.3, 12.4 and 12.5. We see that $\dot{\theta}$, $\dot{u}$, $\dot{\phi}$ and $\dot{\psi}$ are determined by θ, ξ and $\dot{\chi}$. There is no constraint on $\dot{\xi}$.

The configuration manifold Q, as described above, is 7-dimensional, and the constraint distribution D within TQ is 2-dimensional.

12.2 The Group Symmetry

There is a group symmetry of this system associated with the Lie group

$$G = SE(2) \times SO(2)^3.$$

The group action is given using an obvious parameterization for G by

$$
\begin{aligned}
\bar{\theta} &= \theta + \Delta\theta, \\
\bar{u} &= R(\Delta\theta)u + \Delta u \\
\bar{\phi} &= \phi + \Delta\phi, \\
\bar{\psi} &= \psi + \Delta\psi, \\
\bar{\chi} &= \chi + \Delta\chi \qquad \text{and} \\
\bar{\xi} &= \xi + \Delta\theta.
\end{aligned} \tag{12.6}
$$

The quantity α is invariant under this action, and labels the group orbits of the action.

Differentiating with respect to time, we have

$$
\begin{aligned}
\dot{\bar{\theta}} &= \dot{\theta}, \\
\dot{\bar{u}} &= R(\Delta\theta)\dot{u}, \\
\dot{\bar{\phi}} &= \dot{\phi}, \\
\dot{\bar{\psi}} &= \dot{\psi}, \\
\dot{\bar{\chi}} &= \dot{\chi} \qquad \text{and} \\
\dot{\bar{\xi}} &= \dot{\xi}.
\end{aligned}
$$

We see immediately that the Lagrangian (equation 12.1) and the constraints (equations 12.2, 12.3, 12.4 and 12.5) are preserved.

12.3 A Basis Aligned with the Symmetry

We now attempt to apply the theory of chapter 6 to the current group action, given by equations 12.6.

In what follows, we regard TQ as $(\dot{\theta}, \dot{u}, \dot{\phi}, \dot{\psi}, \dot{\chi}, \dot{\xi})$-space.

From the constraints (equations 12.2, 12.3, 12.4 and 12.5), we easily see that $\{\gamma, \kappa_0\}$ is a basis for D, with

$$\gamma = (r\sin\alpha, rR(\theta)\lambda(\alpha), d\cos\alpha - b\sin\alpha, d\cos\alpha + b\sin\alpha, d, r\sin\alpha)$$

and

$$\kappa_0 = (0, 0, 0, 0, 0, 1).$$

It is also easily seen that $\gamma\alpha = 0$ holds, so that γ lies in V, and so in $S = D \cap V$, and that $\kappa_0 \notin V$ holds.

The inner product obtained from equation 12.1 is

$$K\Big((\dot{\theta}, \dot{u}, \dot{\phi}, \dot{\psi}, \dot{\chi}, \dot{\xi}), (\underline{\dot{\theta}}, \underline{\dot{u}}, \underline{\dot{\phi}}, \underline{\dot{\psi}}, \underline{\dot{\chi}}, \underline{\dot{\xi}})\Big) = J\dot{\theta}\underline{\dot{\theta}} + m\dot{u}^T\underline{\dot{u}} + J_W(\dot{\phi}\underline{\dot{\phi}} + \dot{\psi}\underline{\dot{\psi}} + \dot{\chi}\underline{\dot{\chi}}) + J_D\dot{\xi}\underline{\dot{\xi}}. \tag{12.7}$$

This shows that

$$K(\gamma, \gamma) = f(\alpha) \qquad \text{and}$$
$$K(\kappa_0, \gamma) = g(\alpha)$$

hold with

$$f(\alpha) = (J + J_D)r^2\sin^2\alpha + mr^2(d^2\cos^2\alpha + a^2\sin^2\alpha) + J_W(2d^2\cos^2\alpha + 2b^2\sin^2\alpha + d^2)$$

and

$$g(\alpha) = J_D r\sin\alpha.$$

If we define κ by

$$\kappa = f\kappa_0 - g\gamma,$$

then κ lies in $H = D \cap S^\perp$.

Proposition 12.3.1. *The vector fields γ and κ are group invariant.*

Proof. For γ we calculate

$$\begin{aligned}
&\gamma(\overline{\theta}, \overline{u}, \overline{\phi}, \overline{\psi}, \overline{\chi}, \overline{\xi}) \\
&= \gamma(\theta + \Delta\theta, R(\Delta\theta)u + \Delta u, \phi + \Delta\phi, \psi + \Delta\psi, \chi + \Delta\chi, \xi + \Delta\theta) \\
&= (r\sin\alpha, R(\Delta\theta)[rR(\theta)\lambda(\alpha)], \\
&\qquad d\cos\alpha - b\sin\alpha, d\cos\alpha + b\sin\alpha, d, r\sin\alpha) \\
&= (r\sin\alpha, rR(\overline{\theta})\lambda(\alpha), d\cos\alpha - b\sin\alpha, d\cos\alpha + b\sin\alpha, d, 0).
\end{aligned}$$

and for κ_0,

$$\begin{aligned}
&\kappa_0(\bar{\theta}, \bar{u}, \bar{\phi}, \bar{\psi}, \bar{\chi}, \bar{\xi}) \\
&\quad = \kappa_0(\theta + \Delta\theta, R(\Delta\theta)u + \Delta u, \phi + \Delta\phi, \psi + \Delta\psi, \chi + \Delta\chi, \xi + \Delta\theta) \\
&\quad = (0, 0, 0, 0, 0, 1).
\end{aligned}$$

The group invariance of f and g then provides the group invariance of κ. □

12.4 Failure to Satisfy Flatness Conditions

We will now show that this system does not satisfy the flatness conditions.

The flatness conditions, equations 7.2 and 7.3 on page 51, in this case become

$$K(\kappa, [\kappa, \gamma]) \;=\; 0.$$

First we have

$$[\kappa, \gamma] = [f\kappa_0 - g\gamma, \gamma] = f[\kappa_0, \gamma],$$

and then

$$\begin{aligned}
K(\kappa, [\kappa, \gamma]) &= K(f\kappa_0 - g\gamma, f[\kappa_0, \gamma]) \\
&= f\{fK(\kappa_0, [\kappa_0, \gamma]) - gK(\gamma, [\kappa_0, \gamma])\}.
\end{aligned}$$

Our calculation may be greatly simplified by recognizing at this point that

$$\mathcal{L}_{\kappa_0} K = 0$$

holds. To see this, one observes that the coefficients in the inner product of equation 12.7 do not depend on any of the parameters and that $\kappa_0 = \partial/\partial\xi$ holds. We may then calculate

$$\begin{aligned}
g' &= \kappa_0\{K(\kappa_0, \gamma)\} \\
&= (\mathcal{L}_{\kappa_0})(\kappa_0, \gamma) + K([\kappa_0, \kappa_0], \gamma) + K(\kappa_0, [\kappa_0, \gamma]) \\
&= K(\kappa_0, [\kappa_0, \gamma])
\end{aligned}$$

and

$$\begin{aligned} f' &= \kappa_0\{K(\gamma,\gamma)\} \\ &= (\mathcal{L}_{\kappa_0})(\gamma,\gamma) + K([\kappa_0,\gamma],\gamma) + K(\gamma,[\kappa_0,\gamma]) \\ &= 2K(\gamma,[\kappa_0,\gamma]). \end{aligned}$$

For flatness then we need

$$f\left\{fg' - \frac{1}{2}gf'\right\} = K(\kappa,[\kappa,\gamma]) = 0.$$

which leads to

$$g = Cf^{1/2},$$

for some constant C. This is clearly not satisfied.

We point out now that the vector field κ_0 satisfies the conditions on σ given in proposition 5.1.1 on page 40, and so generates a constant of the motion.

We have $\kappa_0 \in D$ by design, and the flow of κ_0 is given by

$$\kappa_{0_s}(\theta,u,\alpha,v,\chi,\xi) = (\theta,u,\alpha,v,\chi,\xi+s),$$

so that $T\kappa_{0_s}$ is given by

$$T\kappa_{0_s}(\dot{\theta},\dot{u},\dot{\alpha},\dot{v},\dot{\chi},\dot{\xi}) = (\dot{\theta},\dot{u},\dot{\alpha},\dot{v},\dot{\chi},\dot{\xi}),$$

and κ_0 fully preserves the Lagrangian.

The constant of the motion generated by κ_0 is $J_D\dot{\xi}$. This is the angular momentum of the front wheel about the vertical axis through its center. It is apparent physically that this should be constant, as the torque on the front wheel about this axis is zero.

Moreover, this constant is group invariant, as it must be, since κ_0 is group invariant (recall equation 7.1 on page 49).

The short coming is that κ_0 is not vertical, and so does not lead to a reduction of the problem using the current theory.

Chapter 13

Chaplygin Sphere

In this section we consider the example of a balanced but dynamically asymmetric sphere rolling without slipping on a horizontal plane. That is, the sphere has its center of mass at its center, but the three principal moments of inertia are (in general) all different. This problem is examined briefly in [3], and was completely integrated by S. Chaplygin (see [10]).

13.1 Formulating the Problem

In this subsection we formulate the Lagrangian and the constraint equations for this system. We will use the notation of section 2.5 in this section, as usual.

We take a fixed set of cartesian axes, with associated orthonormal vectors e_1, e_2 and e_3, with e_3 pointing upwards. We denote the position of the center of mass with respect to these axes by the column vector x.

We also take a set of orthonormal vectors $\{w_1, w_2, w_3\}$ at the center of mass of the sphere and moving with it, and use the orientation matrix

$$W = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}.$$

We denote the angular velocity with respect to the moving axes by ω, so that

$$\dot{W} = WA(\omega)$$

holds.

Let the mass of the sphere be m. We use r to denote the radius of the sphere. Denote the inertia of the body by J. The total kinetic energy of the body is given by

$$K = \frac{1}{2}\omega^T J\omega + \frac{1}{2}m\dot{x}^T\dot{x}.$$

The potential energy of the body is constant, and may be taken to be zero. The Lagrangian for this system is therefore given by

$$L = \frac{1}{2}\omega^T J\omega + \frac{1}{2}m\dot{x}^T\dot{x}. \tag{13.1}$$

The *rolling condition* is that the velocity of the point on the sphere that is in contact with the plane has zero velocity. The constraint therefore is

$$\begin{aligned} 0 &= \dot{x} + A(W\omega)(-re_3) \\ &= \dot{x} + rA(e_3)W\omega. \end{aligned} \tag{13.2}$$

Note that using x and W as coordinates suggests that the configuration space is $SO(3) \times \Re^3$. However, $e_3^T x$ is clearly equal to r, and so the actual configuration manifold is $Q = SO(3) \times \Re^2$.

Also, equation 13.2 suggests that there are three linear velocity constraints. But one of these simply requires that $e_3^T x$ be constant. There are in fact only two linear nonholonomic constraints in TQ.

So the configuration manifold Q is 5-dimensional, and the constraint distribution D within TQ is 3-dimensional.

13.2 The Group Symmetry

In this section we describe a group symmetry of this system associated with the Lie group $G = SE(2)$.

Recall that the Lie group $SE(3)$ is the manifold $SO(3) \times \Re^3$ with the group product given by

$$(H, y)(K, z) = (HK, y + Hz).$$

We will think of $SE(2)$ as the Lie subgroup $\{(R_3(\phi), (\xi, \zeta, 0)) | \phi, \xi, \zeta \in \Re\}$ of $SE(3)$.

To describe the group action, we parameterize G by $\mathbb{R}^3$ using the map

$$(\phi,\xi,\zeta) \mapsto ((R_3(\phi),(\xi,\zeta,0))).$$

The associated group action on $SO(3) \times \mathbb{R}^3$ is given by the map $(W,x) \mapsto (\overline{W},\overline{x})$, with

$$\overline{W} = R_3(\phi)W \tag{13.3}$$

$$\overline{x} = R_3(\phi)x + \xi e_1 + \zeta e_2. \tag{13.4}$$

Since $e_3^T\overline{x} = e_3^T x$ holds, the group action of G on $SO(3) \times \mathbb{R}^3$, induces an action on the configuration manifold Q.

We must show that the Lagrangian,

$$L = \frac{1}{2}\omega^T J\omega + \frac{1}{2}m\dot{x}^T\dot{x},$$

is preserved by this group action.

We immediately have

$$\dot{\overline{x}} = R_3(\phi)\dot{x}.$$

From equation 13.3 we have

$$\begin{aligned}
\overline{W}A(\overline{\omega}) &= \dot{\overline{W}} \\
&= R_3(\phi)\dot{W} \\
&= R_3(\phi)WA(\omega) \\
&= \overline{W}A(\omega),
\end{aligned}$$

so that

$$\overline{\omega} = \omega$$

holds. Thus we obtain $\overline{L} = L$.

We must also show that the constraint,

$$0 = \dot{x} + rA(e_3)W\omega,$$

is preserved. For this we have

$$\begin{aligned}
A(e_3)\overline{W}\overline{\omega} &= A(e_3)R_3(\phi)W\omega \\
&= R_3(\phi)A(e_3)W\omega
\end{aligned}$$

and then

$$\dot{\overline{x}} + rA(e_3)\overline{W}\overline{\omega} = R_3(\phi)\{\dot{x} + rA(e_3)W\omega\}.$$

Since $R_3(\phi)$ is non-singular, we see that the constraint is preserved by the group action.

13.3 A Basis Aligned with the Symmetry

We now apply the theory of chapter 6 to the current group action, given by equation 13.3 and equation 13.4.

To find a basis of vector fields for the distribution V, of vectors tangent to group orbits, we take partial derivatives. We obtain

$$\begin{aligned}
\left.\frac{\partial(\overline{W},\overline{x})}{\partial\xi}\right|_{(0,0,0)} &= (0,\epsilon_1), \\
\left.\frac{\partial(\overline{W},\overline{x})}{\partial\zeta}\right|_{(0,0,0)} &= (0,\epsilon_2) \qquad \text{and} \\
\left.\frac{\partial(\overline{W},\overline{x})}{\partial o}\right|_{(0,0,0)} &= (A(\epsilon_3)W, A(\epsilon_3)x) \\
&= (WA(W^{T}\epsilon_3), A(\epsilon_3)x).
\end{aligned}$$

We may map these to $(\omega, \dot{x})$-space to obtain instead

$$\begin{aligned}
\frac{\partial}{\partial\xi} &\mapsto (0,\epsilon_1), \\
\frac{\partial}{\partial\zeta} &\mapsto (0,\epsilon_2) \qquad \text{and} \\
\frac{\partial}{\partial o} &\mapsto (W^{T}\epsilon_3, A(\epsilon_3)x).
\end{aligned}$$

Each of these vector fields will necessarily take values in TQ when evaluated on Q. These vector fields are also independent everywhere.

To find vector fields which span the constraint distribution D, in $(\omega, \dot{x})$-space, we simply substitute values for ω in the constraint.

$$0 = \dot{x} + rA(\epsilon_3)W\omega.$$

Our choices for ω are made with proposition 8.5.1 on page 69 in mind.

For $\omega = W^{T}\epsilon_3$ we obtain

$$\dot{x} = -rA(\epsilon_3)\epsilon_3 = 0.$$

For $\omega = \epsilon_i$ we obtain

$$\dot{x} = -rA(\epsilon_3)W\epsilon_i.$$

Thus $D = \text{span}\{\overset{0}{\kappa}_1, \overset{0}{\kappa}_2, \overset{0}{\kappa}_3\}$ and $\gamma \in D$ hold, with

$$\begin{aligned} \gamma &= (W^T e_3, 0) \qquad \text{and} \\ \overset{0}{\kappa}_i &= (e_i, -rA(e_3)We_i). \end{aligned}$$

Each of these vector fields will also necessarily take values in TQ when evaluated on Q.

Viewing these vector fields as differentiation operators, we have

$$\begin{aligned} (13.5) \qquad \gamma(W, x) &= (WA(W^T e_3), 0) = (A(e_3)W, 0) \quad \text{and} \\ \overset{0}{\kappa}_i\,(W, x) &= (WA(e_i), -rA(e_3)We_i). \end{aligned}$$

Being somewhat loose in our notation, we see now that

$$\gamma = \frac{\partial}{\partial \phi} + \{e_2^T x\}\frac{\partial}{\partial \xi} - \{e_1^T x\}\frac{\partial}{\partial \zeta}$$

holds, so that γ lies in $S = D \cap V$.

We orthogonally project each vector field $\overset{0}{\kappa}_i$ onto the subspace $D \cap \gamma^\perp$ by

$$\kappa_i = K(\gamma, \gamma)\,\overset{0}{\kappa}_i - K\left(\gamma, \overset{0}{\kappa}_i\right)\gamma.$$

Then $\{\kappa_1, \kappa_2, \kappa_3\}$ spans $D \cap \gamma^\perp$, so that $\{\gamma, \kappa_1, \kappa_2, \kappa_3\}$ spans D. Moreover we have

$$\kappa_i \in V \Rightarrow \overset{0}{\kappa}_i \in V \Rightarrow e_i = W^T e_3 \Rightarrow \overset{0}{\kappa}_i = \gamma \Rightarrow \kappa_i = 0.$$

It follows that $\{\gamma\}$ spans S, and $\{\kappa_1, \kappa_2, \kappa_3\}$ spans $H = D \cap S^\perp$.

At any point then, for some choice of i and j, $\{\gamma, \kappa_i, \kappa_j\}$ is a basis aligned with the symmetry.

Proposition 13.3.1. *The vector fields γ, κ_1, κ_2 and κ_3 are group invariant.*

Proof. Since K is group invariant, it is sufficient to show that γ, $\overset{0}{\kappa}_1$, $\overset{0}{\kappa}_2$ and $\overset{0}{\kappa}_3$ are group invariant. In light of proposition 8.5.1, it is sufficient to compare the effect of each vector on $\overline{W}$, with its effect on W.

For γ we have

$$\begin{aligned} \gamma\overline{W} &= \gamma\{R_3(\phi)W\} \\ &= R_3(\phi)A(e_3)W \\ &= A(e_3)R_3(\phi)W \\ &= A(e_3)\overline{W}. \end{aligned}$$

For $\overset{0}{\kappa}_i$ we have

$$\begin{aligned}\overset{0}{\kappa}_i\, \overline{W} &= \overset{0}{\kappa}_i\, \{R_3(\phi)W\} \\ &= R_3(\phi)WA(\epsilon_i) \\ &= \overline{W}A(\epsilon_i).\end{aligned}$$

□

13.4 The Flatness Conditions

Lemma 13.4.1. *In $(\omega, \dot{x})$-space we have*

$$[\gamma, \overset{0}{\kappa}_i] = (0, -rA(\epsilon_3)A(\epsilon_3)W\epsilon_i).$$

Proof. We calculate

$$\begin{aligned}[\gamma, \overset{0}{\kappa}_i]W &= \gamma\{WA(\epsilon_i)\} - \overset{0}{\kappa}_i\, \{A(\epsilon_3)W\} \\ &= A(\epsilon_3)WA(\epsilon_i) - A(\epsilon_3)WA(\epsilon_i) \\ &= 0\end{aligned}$$

and

$$\begin{aligned}[\gamma, \overset{0}{\kappa}_i]x &= \gamma\{-rA(\epsilon_3)W\epsilon_i\} - \overset{0}{\kappa}_i\, \{0\} \\ &= -rA(\epsilon_3)A(\epsilon_3)W\epsilon_i.\end{aligned}$$

□

The kinetic energy for this system, once again is

$$K = \frac{1}{2}\omega^T J\omega + \frac{1}{2}m\dot{x}^T\dot{x}.$$

The associated inner product, in $(\omega, \dot{x})$-space, is given by

$$K((\lambda, u), (\mu, v)) = \lambda^T J\mu + mu^T v.$$

Proposition 13.4.1. *The basis $\{\gamma, \kappa_i, \kappa_j\}$ satisfies the flatness conditions, equations 7.2 and 7.3 on page 51, which in this case become*

$$K([\gamma, \kappa_i], \kappa_j) + K([\gamma, \kappa_j], \kappa_i) = 0.$$

Proof. First we calculate

$$[\gamma, \kappa_i] = \left[\gamma, K(\gamma,\gamma)\overset{0}{\kappa}_i - K\left(\gamma, \overset{0}{\kappa}_i\right)\gamma\right] = K(\gamma,\gamma)[\gamma, \overset{0}{\kappa}_i].$$

and then

$$\begin{aligned} K([\gamma,\kappa_i],\kappa_j) &= K\left(K(\gamma,\gamma)[\gamma,\overset{0}{\kappa}_i], K(\gamma,\gamma)\overset{0}{\kappa}_j - K\left(\gamma,\overset{0}{\kappa}_j\right)\gamma\right) \\ &= \{K(\gamma,\gamma)\}^2 K\left([\gamma,\overset{0}{\kappa}_i],\overset{0}{\kappa}_j\right) \\ &\qquad - K(\gamma,\gamma)\, K\left(\gamma,\overset{0}{\kappa}_j\right) K\left([\gamma,\overset{0}{\kappa}_i],\gamma\right). \end{aligned}$$

We need this last expression to be anti-symmetric in i and j. But lemma 13.4.1 immediately gives $K\left([\gamma,\overset{0}{\kappa}_i],\gamma\right) = 0$ and

$$\begin{aligned} K\left([\gamma,\overset{0}{\kappa}_i],\overset{0}{\kappa}_j\right) &= m\{-rA(\epsilon_3)A(\epsilon_3)W\epsilon_i\}^T\{-rA(\epsilon_3)W\epsilon_j\} \\ &= -mr^2\epsilon_j^T W^T A(\epsilon_3)A(\epsilon_3)A(\epsilon_3)W\epsilon_i. \end{aligned}$$

This last expression is indeed anti-symmetric. □

13.5 The Adjoint Equation

We will use suffixes for the components of Ω, in order to have consistent notation in this section.

Proposition 13.5.1. *The adjoint equation for the Chaplygin sphere becomes*

$$\kappa_j\{K(\gamma,\gamma)\Omega_1\} = 0,$$

and can therefore be solved by

$$\Omega_1 = \frac{C}{K(\gamma,\gamma)},$$

where C is a constant of integration.

Proof. For this problem, equation 7.11 on page 57 becomes

$$\kappa_j\{K(\gamma,\gamma)\Omega_1\} + K([\gamma,\kappa_j],\gamma)\Omega_1 = 0.$$

But, as in the proposition on flatness, we have

$$[\gamma, \kappa_j] = K(\gamma, \gamma)[\gamma, \overset{0}{\kappa}_j].$$

and then

$$K([\gamma, \kappa_j], \gamma) = K(\gamma, \gamma) K\left([\gamma, \overset{0}{\kappa}_j], \gamma\right) = 0.$$

using lemma 13.4.1. □

In this case, Q/G is simply the surface of a sphere, the manifold S^2. Since the dimension of S^2 is 2, we conclude by corollary 8.4.1 on page 68, and the comments that follow corollary 8.4.1, that the reduced system on S^2 is Hamiltonian.

We can interpret the reduced system physically. First, if we think of S^2 as $\{y \in \mathbb{R}^3 | y^T y = 1\}$, then the projection map $\Lambda : Q = SO(3) \times \mathbb{R}^2 \to S^2$ is given by $(W, u) \mapsto W^T e_3$ That is, $W^T e_3$ labels the group orbits. But recall that if a material point in the rolling sphere is labelled with respect to the moving axes by y, then with respect to the fixed axes it is labelled by $z = x + Wy$. If we set z to the position of the point of contact, then $x - z = re_3$ holds, and the corresponding value for y is $-rW^T e_3$. So we may think of S^2 as the surface of the rolling sphere, and then the trajectory in the reduced system will be the path traced out on the surface of the rolling sphere, by the point of contact.

13.6 Constants of the Motion

The constant C in the previous subsection has an interpretation in terms of the angular momentum of the sphere. We ask the reader to recall equation 2.10 on page 18, and the notation in use in that subsection. There, the angular momentum about the origin was the quantity $P = mA(x)\dot{x} + WJ\omega$. The expression $WJ\omega$ was called the angular momentum of the body.

From the general theory, we know that the tangent vector to the trajectory, $\tau(t)$, will be a linear combination of γ, κ_i and κ_j. So we have

$$\begin{aligned}
e_3^T\{WJ\omega(\tau)\} &= (W^T e_3)^T J\omega(\Omega_1\gamma + v_i\kappa_i + v_j\kappa_j) \\
&= \omega(\gamma)^T J\omega(\Omega_1\gamma + v_i\kappa_i + v_j\kappa_j) \\
&= K(\gamma, \Omega_1\gamma + v_i\kappa_i + v_j\kappa_j) \\
&= K(\gamma, \gamma)\Omega_1 \\
&= C.
\end{aligned}$$

Thus C is the component of the angular momentum of the body, along the vertical axis.

We also point out that the vector field γ satisfies the conditions on σ given in proposition 5.1.1 on page 40, and so generates a constant of the motion. This constant is in fact C.

To see this, we have $\gamma \in D$ by design, and the flow of γ is given by

$$\gamma_s(W, x) = (\overline{W}, \overline{x})$$

with $\overline{W} = R_3(s)W$ and $\overline{x} = x$. Differentiating this gives $\overline{\omega} = \omega$ and $\dot{\overline{x}} = \dot{x}$, in the ususal way, so that $T\gamma_s$ is given in $(\omega, \dot{x})$-space by

$$T\gamma_s(\omega, \dot{x}) = (\omega, \dot{x}),$$

which is the identity map. So γ fully preserves the Lagrangian,

$$L = \frac{1}{2}\omega^T J\omega + \frac{1}{2}m\dot{x}^T\dot{x}.$$

The constant of the motion generated by γ then must be

$$K(\tau, \gamma) = \omega(\tau)^T J\{W^T e_3\} = e_3^T WJ\omega(\tau) = C.$$

In fact, C is one of three related constants. Let σ be the vector field given in $(\omega, \dot{x})$-space by

$$\sigma = (W^T e_i, rA(e_i)e_3).$$

For $i = 3$ this is just γ. We have $\sigma \in D$, and the flow of σ is given by

$$\sigma_s(W, x) = (\overline{W}, \overline{x})$$

with $\overline{W} = R_i(s)W$ and $\overline{x} = R(srA(e_i)e_3)x$. Differentiating this gives $\overline{\omega} = \omega$ and $\dot{\overline{x}} = R(srA(e_i)e_3)\dot{x}$, so that $T\gamma_s$ is given in $(\omega, \dot{x})$-space by

$$T\gamma_s(\omega, \dot{x}) = (\omega, R(srA(e_i)e_3)\dot{x}).$$

We see that this fully preserves the Lagrangian. The constant of the motion generated by σ then must be

$$\begin{aligned} C_i &= K(\tau, \sigma) \\ &= \omega(\tau)^T J\{W^T e_i\} + m\dot{x}(\tau)^T\{rA(e_i)e_3\} \\ &= e_i^T\{WJ\omega(\tau) + mA(re_3)\dot{x}(\tau)\}. \end{aligned}$$

This last expression is very nearly the same as the expression $P = mA(x)\dot{x} + W J\omega$ for total angular momentum. We could say that the constant is the component along the ϵ_i-direction. of the total angular momentum of the body about the point of contact. To do so however ignores that the fact that the point of contact is not fixed (in any Galilean coordinate system).

The additional two constants found in this way are not group invariant. We find

$$\begin{aligned} \overline{C}_i &= \epsilon_i^T\{\overline{W}J\overline{\omega} + mA(r\epsilon_3)\dot{\overline{x}}\} \\ &= \epsilon_i^T\{R_3(o)WJ\omega + mA(r\epsilon_3)R_3(o)\dot{x}\} \\ &= \epsilon_i^T R_3(o)\{WJ\omega + mA(r\epsilon_3)\dot{x}\}. \end{aligned}$$

which is only equal to C_i for $i = 3$. However

$$\sum_i (\overline{C}_i)^2 = \{WJ\omega + mA(r\epsilon_3)\dot{x}\}^T\{WJ\omega + mA(r\epsilon_3)\dot{x}\} = \sum_i (C_i)^2$$

does hold.

The energy is of course also a constant of the motion. There are therefore two constants of the motion. quadratic in the velocities. In [10]. Chaplygin used these constants in conjunction with elliptic coordinates on S^2 to integrate this system.

Bibliography

[1] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. Addison-Wesley, sixth edition, 1987.

[2] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer-Verlag, second edition, 1989.

[3] V. I. Arnold, editor. *Dynamical Systems III*. Springer-Verlag, 1993.

[4] L. Bates, H. Graumann, and C. MacDonnell. Examples of gauge conservation laws in nonholonomic systems. *Preprint*, 1995.

[5] L. Bates and J. Śniatycki. Nonholonomic reduction. *Reports on Mathematical Physics*, 32:99–115, 1993.

[6] R. L. Bishop and S. I. Goldberg. *Tensor Analysis on Manifolds*. Dover, 1980.

[7] A. Bloch, P. Krishnaprasad, J. Marsden, and R. Murray. Nonholonomic mechanical systems with symmetry. Technical report, California Institute of Technology, 1994.

[8] L. Boltzmann. *Wien. Sitzungsberichte*, 1902.

[9] C. M. Braams. On the influence of friction on the motion of a top. *Physica*, 18:503–514, 1952.

[10] S. Chaplygin. *Investigations of the Dynamics of Non-Holonomic Systems*. Gostekhizdat, Moscow-Leningrad, 1949. Russian.

[11] R. J. Cohen. The tippe top revisited. *American Journal of Physics*, 45:12–17, 1977.

[12] R. Cushman, J. Hermans, and D. Kemppainen. The rolling disc. *Progress in Nonlinear Differential Equations and Their Applications*, 19:21–60, 1996.

[13] S. Ebenfeld and F. Scheck. A new analysis of the tippe top: Asymptotic states and liapunov stability. *Preprint*, 1995.

[14] N. M. Ferrers. *Quart. Journ. Math.*, XII:1, 1871.

[15] A. D. Fokker. The rising top, experimental evidence and theory. *Physica*, 8:591–514, 1941.

[16] E. G. Gallop. On the rise of a spinning top. *Cambridge Philosophical Transactions*, XIX:356–373, 1903.

[17] H. Goldstein. *Foundations of Mechanics*. Addison-Wesley, 1950.

[18] G. Hamel. *Zeitschrift für Math. u. Phys.*, 1904.

[19] J. Hermans. A symmetric sphere rolling on a surface. *Nonlinearity*, 8:493–515, 1995.

[20] N. M. Hugenholtz. On tops rising by friction. *Physica*, 18:515–527, 1952.

[21] J. D. Jackson. *Classical Electrodynamics*. John Wiley and Sons, second edition, 1975.

[22] J. Jellett. *Theory of Friction*. Hodges, Foster, Dublin, 1872.

[23] T. R. Kane and D. A. Levinson. A realistic solution to the symmetric top problem. *Journal of Applied Mechanics*, 45:903–909, 1978.

[24] D. Kemppainen. The geometry of the rolling disk. Master's thesis, The University of Calgary, 1994.

[25] J. Koiller. Reduction of some classical non-holonomic systems with symmetry. *Arch. Rational Mach. Anal.*, 118:113–148, 1992.

[26] C. Neumann. *Leipzig Berichte*, XL:22, 1888.

[27] A. C. Or. The dynamics of the tippe top. *SIAM J. Appl. Math.*, 54:597–609, 1994.

[28] D. G. Parkyn. The rising of tops with rounded pegs. *Physica*, 24:313–330, 1958.

[29] L. Pars. *Analytical Dynamics*. Heinemann, New York, 1965.

[30] W. A. Pliskin. The tippe top (topsy-turvy top). *American Journal of Physics*, 22:28–32, 1954.

[31] E. Routh. *The Advanced Part of a Treatise on the Dynamics of a System of Rigid Bodies*. Macmillan, London, sixth edition, 1905. Dover edition - 1955.

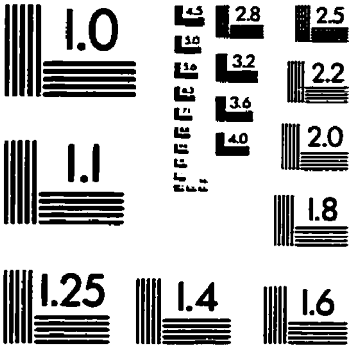
[32] A. Vierkandt. *Monatshefte für Math. u. Phys.*, III:31, 1892.

[33] F. Warner. *Foundations of Differentiable Manifolds and Lie Groups*. Scott, Foresman, 1971.

[34] E. T. Whittaker. *Analytical Dynamics*. Cambridge University Press, fourth edition, 1937.

Index

TEST TARGET (QA-3)



150mm

6″

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989