RELATIVE EQUILIBRIA OF COUPLED UNDERWATER VEHICLES

A Thesis Submitted to the College of
Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in the Department of Mathematics & Statistics
The University of Saskatchewan
Saskatoon

by
Natalia Pavlovna Fomenko

Abstract

The dynamics of a single underwater vehicle in an ideal irrotational fluid may be modeled by a Lagrangian system with configuration space the Euclidean group. If hydrodynamic coupling is ignored then two coupled vehicles may be modeled by the direct product of two single-vehicle systems. We consider this system in the case that the vehicles are coupled mechanically, with an ideal spherically symmetric joint, finding all of the relative equilibria. We demonstrate that there are relative equilibria in certain novel momentum-generator regimes identified by Patrick et.al. "Stability of Poisson equilibria and Hamiltonian relative equilibria by energy methods", *Arch. Rational Mech. Anal.*, 174:301–344, 2004.

Acknowledgments

This thesis was completed with the encouragement, support, and collaboration of my supervisor, Professor Patrick.

Contents

Introduction

A relative equilibrium is a special motion of a symmetric system which is consistent with its symmetry. The technical definition: for a system with symmetry group G acting on phase space P, a relative equilibrium is a point $p_e \in P$ and a *generator* $\xi_e \in \mathfrak{g}$ such that $t \mapsto (\exp \xi_e t)p_e$ is an evolution of the system. In a mechanical system, symmetries are associated to conserved quantities called momenta, and these are realized as a momentum mapping, which is a map from phase space to the dual $\mathfrak{g}^*$ of the Lie algebra of the symmetry group. So every relative equilibrium p_e has a momentum $\mu_e = J(p_e)$, and associated to any relative equilibrium is its momentum-generator pair $(\mu_e, \xi_e) \in \mathfrak{g}^* \times \mathfrak{g}$.

Chapter 1 of this thesis provides a self contained exposition of manifolds and Lie groups. For simplicity, a manifold is defined as smooth subset of $\mathbb{R}^n$, and a Lie group is defined as smooth subset of the group of invertible $p \times p$ matrices. Chapter 2 of this thesis provides a self contained exposition of mechanical systems with symmetry, using the Lagrangian variational formalism. Full technical definitions of the constructs *Lie group, Lie algebra, Lagrangian system, evolution, momentum, relative equilibrium, and generator*, can be found in these two preliminary chapters. This entire context is illustrated in Chapter 3 using the well known mechanical system of a free single rigid body, as can be found in, for example, in [1].

An equilibria of a mechanical system is a point p_e of phase space such the constant curve at p_e is an evolution. Equilibria can be stable or unstable. A standard example of a system with both stable and unstable equilibria is the simple pendulum. When the pendulum rests with its bob hanging down, then a sufficiently small perturbation of it results in a motion that remains arbitrarily close to this resting motion. On the other hand, if the pendulum is resting perfectly with its bob suspended straight up, then any small perturbation will cause the pendulum to fall. The stability of an equilibrium can be determined from the system's energy: if the second derivative of the energy is definite at the equilibrium, then the energy levels of the system near the equilibrium surround the equilibrium. By conservation of energy, this confines evolutions starting near the equilibrium to be forever near the equilibrium.

The stability of relative equilibria is a more involved matter. *Testing for stability is an involved procedure* [6]. For a compact group, one tests for definiteness of the function $H - \langle J, \xi_e \rangle$ on $\ker dJ(p_e)/\mathfrak{g}_{\mu_e}$, which means on any complement to $\mathfrak{g}_{\mu_e}$ in $\ker dJ(p_e)$. *The system will drift* [8, 9, 10, 11] in the direction of the symmetry group: it drifts along the coadjoint isotropy group G_{μ_e} of the momentum μ_e.

The theory of the stability of relative equilibria in the case of non-compact groups, such as the Euclidean group $SE(3)$, has recently been developed [13]. For some momentum-generator pairs, stability by energy confinement is impossible. Such relative equilibria are called *wild*, and determining their stability is very delicate. For example, the stability of pure rotation relative equilibria of a single underwater vehicle has been determined by KAM theory in [12]. A relative equilibria is called *tame* if it is not wild. For tame relative equilibria, the subspace on which definiteness must be established—the *test subspace*—varies by the momentum-generator pair and is sometimes strictly larger than $\ker dJ(p_e)/\mathfrak{g}_{\mu_e}$. Finally, the kind of stability varies by the momentum-generator pair as well: some of these pairs have an *exotic stability type* which is strictly larger that G_{μ_e}.

One of the first works to investigate the stability of relative equilibria in the context of noncompact symmetry was [5], where Leonard and Marsden considered a single vehicle immersed in an ideal, irrotational fluid. The relative equilibria of this system are uniformly translating and rotating motions. In Chapters 4 and 5 of this thesis we calculate the relative equilibria the physical system of two underwater vehicles coupled with spherically symmetric joint. The vehicles will be axially symmetric and the symmetry group of the system will be $SE(3) \times SO(2)^2$.

From symmetry considerations alone, Patrick, Roberts, and Wulff [13] identify interesting momentum-generator regimes for $SE(3) \times SO(2)^2$, which have not to date been realized in the literature in an analysis of relative equilibria of a physical system. Those momentum-generator regimes are as follows: The part of the momentum associated to $SE(3)$ can be written (μ^r, μ^a), where μ^r is a rotation part and μ^a is a translation part. Similarly the generator associated to $SE(3)$ is (ξ^r, ξ^a).

1. $\mu_e^a = 0$, $\xi_e^r = 0$, $\mu_e^r \neq 0$. These relative equilibria do not translate, have no overall rotation, but have a angular momentum due the spinning of the separate vehicles on their axes. They are tame with exotic stability type. Such relative equilibria cannot be realized in the system of a single underwater vehicle because with no multiple rotating parts, a single underwater vehicle has zero rotational momentum when $\xi_e^r = 0$.

2. $\mu_e^a = 0$, $\xi_e^r \neq 0$, $\mu_e^r = 0$. These relative equilibria do not translate, have an overall rotation, but, because of the multiple rotating parts, they have zero net rotational angular momentum. These are wild relative equilibria, and so their stability cannot be established by using energy and momentum as Lyapunov functions. However, given stability, the drift that is expected to occur will be along the whole symmetry group, and so these are interesting relative equilibria. Drift such as this was investigated for coupled rigid bodies in [9, 10, 11].

At the end of Chapter 5, we determine that there are relative equilibria of the system of coupled underwater vehicles in these two momentum-generator regimes.

Chapter 1

Preliminaries

In this thesis the Einstein sum convention is in effect, so that a product where an index occurs twice, once as a superscript and once as a subscript, means the sum over that index pair. Thus, $x^i y_i = x^1 y_1 + x^2 y_2 + \cdots$.

Unless explicitly stated, all vectors in $\mathbb{R}^n$ are considered as column vectors, even through for typographical reasons they may be written as row vectors. The transpose of a matrix A will be denoted A^t.

Let A be a $n \times m$ real matrix. Suppose the linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ defined by $x \mapsto Ax$ is surjective (i.e. the columns of A span $\mathbb{R}^n$). Then the rows of A are a basis of the orthogonal complement of $\ker A$. Indeed, if the rows of A are denoted A^i and $x \in \ker A$, then Ax is the column vector of dot products $A^i \cdot x$, so $A^i \cdot x = 0$ for all i, hence $A^i \in \ker A^\perp$ for all i. If there is a nontrivial linear combination of the A^i which is zero, say $c_i A^i = 0$. Then $c_i A^i_j x^j = 0$ for all x. This contradicts $x \mapsto Ax$ surjective because it implies that Ax takes values in a proper subspace. Thus the n rows of A^i are a linearly independent set of vectors in $\ker A^\perp$. These are a basis since $\dim \ker A^\perp = m - \dim \ker A = m - (m - n) = n$. Since these rows are linearly independent, the transpose matrix A^t is such that the map $\alpha \mapsto A^t \alpha$ is injective.

We will often have use for the following formula from elementary calculus. Suppose t occurs in an expression y in multiple places and we wish to calculate the derivative dy/dt of y with respect to t at $t = a$. We may sequentially set all but one of the t equal to a and differentiate at $t = a$, and sum those derivatives to obtain

$\dfrac{dy}{dt}(a)$. Indeed, this is the situation where $y = f(t, t, \ldots, t)$, so $y = f(z)$ where $z = (t, t, \ldots, t)$. Then by the chain rule

$$\frac{dy}{dt}(a) = \frac{\partial y}{\partial z^i}(a, a, \ldots, a)\frac{dz^i}{dt}(a) = \sum_{i=1} \frac{\partial y}{\partial z^i}(a, a, \ldots, a),$$

as required.

1.1 Manifolds

Let $f \colon U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be a C^1 map. We recall that the *derivative of* f is the map $x \mapsto Df(x)$ where $Df(x)$ is the linear map

$$Df(x)\delta x = \left.\frac{d}{dt}\right|_{t=0} f(x + t\delta x).$$

Equivalently, $Df(x)$ is the matrix

$$Df(x) = \begin{bmatrix} \dfrac{\partial f^1}{\partial x^1} & \cdots & \dfrac{\partial f^1}{\partial x^m} \\ \vdots & & \vdots \\ \dfrac{\partial f^n}{\partial x^1} & \cdots & \dfrac{\partial f^n}{\partial x^m} \end{bmatrix}.$$

If $f \colon U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is any map and $a \in \mathbb{R}^n$, then the *a level set* of f, denoted $f^{-1}(a)$ is the set $f^{-1}(a) = \{x \in U \mid f(x) = a\}$.

A C^1 map $f \colon U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is a *submersion* if, for all $x \in U$, the linear map $Df(x)$ is onto. In this thesis, for simplicity, a *manifold* means a subset of $\mathbb{R}^m$ which is the level set of a given C^∞ submersion. If $M \subseteq \mathbb{R}^m$ and $N \subseteq \mathbb{R}^n$ are manifolds, a map $f \colon M \to N$ is called C^k, $k \geq 0$, if there is a C^k map $\hat{f} \colon U \to \mathbb{R}^n$ such that $U \subseteq \mathbb{R}^m$ is open, $M \subseteq U$, and $\hat{f}|M = f$. A C^k *diffeomorphism*, $k \geq 1$, is a C^k map with a C^k inverse.

The following theorem provides assurance that, near any point of a manifold, there are coordinates in the ambient space in which the manifold becomes and open subset of a subspace. The proof is provided here for completeness, but it can also be found in any text on differentiable manifolds in the place that the level set of a submersion is shown to be a submanifold.

Theorem 1 *Let $M = f^{-1}(a)$, where $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is a submersion, and let $x_0 \in M$. Then there is a C^∞ diffeomorphism $\psi\colon U \subseteq \mathbb{R}^m \to V \subseteq \mathbb{R}^{m-n} \times \mathbb{R}^n$ such that $x_0 \in U$ and*

$$x \in M \cap U \Leftrightarrow x \in U \text{ and } \psi(x) \in \mathbb{R}^{m-n} \times \{0\}$$

Proof. The proof is trivial if $n = 0$. Otherwise, since $Df(x_0)$ is onto, its n rows are a basis of the *orthogonal complement* of $\ker Df(x_0)$. Choose an $(m-n) \times m$ dimensional matrix B which has rows which are a basis of $\ker Df(x_0)$, and define h by

$$h(x) = \big(B(x - x_0), f(x) - a\big).$$

Then h is C^∞ and

$$Dh(x_0) = \begin{bmatrix} B \\ \\ Df(x_0) \end{bmatrix}.$$

$Dh(x_0)$ is nonsingular and $h(x_0) = (0, 0)$, so the inverse function theorem implies that there are neighborhoods U of x_0 and V of 0 such that $h|U\colon U \to V$ is a diffeomorphism. Define $\psi \equiv h|U$. Then

$$x \in U \cap M \Leftrightarrow x \in U \text{ and } f(x) = a \Leftrightarrow x \in U \text{ and } \psi(x) \in \mathbb{R}^{m-n} \times \{0\},$$

as required. ∎

Remark 2 Given *any* open subset of $\mathbb{R}^m$, and defining $f\colon U \to \{0\}$ by $f(x) = 0$, one sees that any open subset of $\mathbb{R}^m$ is a manifold in the sense of this thesis.

1.2 Tangent bundles

Let $f\colon U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be a C^∞ submersion, so that $M = f^{-1}(a)$ is a manifold. The *tangent bundle TM of M* is the $(a, 0)$ level set of the mapping $Tf\colon U \times \mathbb{R}^m \to \mathbb{R}^{2n}$ defined by $Tf(x, \dot{x}) = \big(f(x), Df(x)\dot{x}\big)$ i.e.

$$TM = \{(x, \dot{x}) \mid f(x) = a, Df(x)\dot{x} = 0\}.$$

By definition, $TM \subseteq R^{2m}$. The *second derivative* of f is the map $x \mapsto D^2 f(x)$ where $D^2 f(x)$ is the symmetric bilinear map

$$D^2 f(x)(\delta x_1, \delta x_2) = \frac{\partial}{\partial s}\bigg|_{s=0} \frac{\partial}{\partial t}\bigg|_{t=0} f(x + s\delta x_1 + t\delta x_2).$$

The derivative of the mapping Tf is

$$\begin{aligned}
(\delta x, \delta\dot{x}) \mapsto \frac{d}{dt}\Big|_{t=0} & Tf(x+t\delta x, \dot{x}+t\delta\dot{x}) \\
&= \frac{d}{dt}\Big|_{t=0} \big(f(x+t\delta x), Df(x+t\delta x)(\dot{x}+t\delta\dot{x})\big) \\
&= \left(\frac{d}{dt}\Big|_{t=0} f(x+t\delta x), \frac{d}{dt}\Big|_{t=0} Df(x+t\delta x)\dot{x} + \frac{d}{dt}\Big|_{t=0} Df(x)(\dot{x}+t\delta\dot{x})\right) \\
&= \left(\frac{d}{dt}\Big|_{t=0} f(x+t\delta x), \frac{\partial}{\partial t}\Big|_{t=0} \frac{\partial}{\partial s}\Big|_{s=0} f(x+t\delta x+s\dot{x}) + Df(x)(\delta\dot{x})\right) \\
&= \big(Df(x)\delta x, D^2f(x)(\delta x, \delta\dot{x}) + Df(x)\delta\dot{x}\big).
\end{aligned}$$

By separately setting $\delta\dot{x} = 0$ and $\delta x = 0$, one obtains every element of the form $\big(Df(x)\delta x, 0\big)$ and $\big(0, Df(x)\delta\dot{x}\big)$, respectively. So Tf is a submersion, since $Df(x)$ is onto, and thus TM is a manifold in the sense of this thesis.

The map $\tau_M : TM \to M$ defined by $\tau_M(x, \dot{x}) = x$ is called the *canonical projection* of TM. τ_M is C^∞ since it is the restriction to TM of the map from $\mathbb{R}^{2m} \to \mathbb{R}^m$ defined by $(x, \dot{x}) \mapsto x$, which is C^∞. The set $\tau_M^{-1}(x)$ is called the *fiber of TM over* x and is also denoted by T_xM. TM is the disjoint union of its fibers, and each fiber is a vector space with the operations

$$a(x, \dot{x}) = (x, a\dot{x}), \quad (x, \dot{x}_1) + (x, \dot{x}_2) = (x, \dot{x}_1 + \dot{x}_2).$$

Remark 3 Let $x_0 \in M$ and $x(t)$ be a C^1 curve with $x(t) \in M$ for all t. Then $f\big(x(t)\big) = a$ for all t. Differentiating at $t = 0$ gives $Df(x_0)x'(0) = 0$, so $\big(x, x'(0)\big) \in T_{x_0}M$. Conversely, let $(x, \delta x) \in T_{x_0}M$. Let h be the function defined in the proof of Theorem 1. By the definition of T_xM, $Df(x_0)\delta x = 0$, so $Dh(x_0)\delta x \in \mathbb{R}^{m-n} \times \{0\}$, and the curve $x(t) = \psi^{-1}\big(tDh(x_0)\delta x\big)$, which is defined for t small enough, is in M. Then $x'(0) = D\psi^{-1}(0)Dh(x_0)\delta x = \delta x$, so δx is the derivative of a curve in M. Thus

$$T_{x_0}M = \Big\{ \big(x_0, x'(0)\big) \;\Big|\; x(0) = x_0 \text{ and } x(t) \in M \Big\},$$

i.e. the tangent bindle of a manifold is the set of derivatives of curves in the manifold, with the base point added.

1.3 Vector fields and integral curves

We continue with $M = f^{-1}(a)$, where $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is a submersion. A C^∞ *vector field* on M is a C^∞ map $X\colon M \to TM$ such that $\big(x, X(x)\big) \in T_xM$ for all $x \in M$. An *integral curve of* X is a C^1 curve $x(t) \in M$, $t \in (a,b)$ such that $x'(t) = X\big(x(t)\big)$, where $'$ denotes the derivative with respect to t. If $x_0 \in M$ and $t \in \mathbb{R}$, then an integral curve $x(t)$ *starts at* x_0 *at time* t_0 if $x(t_0) = x_0$. We recall the main existence and uniqueness theorem of ordinary differential equations (for more details, see [2]).

Theorem 4 *Let $X\colon U \subseteq \mathbb{R}^m \to \mathbb{R}^m$ be a C^k vector field, $k \geq 1$, and let $x_0 \in U$ and $t_0 \in \mathbb{R}$. Then there is a unique integral curve $x(t)$, $t \in (a,b)$, of X starting at x_0 at time t_0, and such that every integral curve of X starting at x_0 at time t_0 is a restriction of $x(t)$ to some open interval contained in (a,b).*

In components, the condition that $x(t)$ is an integral curve starting at x_0 at time t_0 is exactly

$$\frac{dx^i}{dt} = X^i\big(x(t)\big), \qquad x(t_0) = x_0.$$

This is the generic initial value problem of ordinary differential equations. Strictly speaking, integral curves are never unique, since given any integral curve there is always a second one by restricting the domain of the first. Theorem 4 avoids this inconvenience by asserting the existence of a unique *maximal integral curve* i.e. a unique curve which is defined for as long a time as is possible in the sense that every integral curve is a restriction of the maximal one. The interval of existence of the maximal integral curve is dictated by the vector field X and the initial conditions x_0 and t_0. This interval must be determined as a part of the problem of solving the differential equation $x'(t) = X\big(x(t)\big)$.

Theorem 4 is stated for vector fields defined on open subsets of $\mathbb{R}^m$. It is also true for vector fields on manifolds, and completeness we provide the following theorem.

Theorem 5 *Let X be a C^k vector field on M, $k \geq 1$, and let $x_0 \in U$ and $t_0 \in \mathbb{R}$. Then there is a unique maximal integral curve of X starting at x_0 at time t_0.*

Proof. Define the map $P\colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ so that $P(x, \dot{x})$ is the orthogonal projection of $\dot{x}$ onto $\ker Df(x)$. P is C^∞. Indeed, the orthogonal complement of $\ker Df(x)$

has basis the rows of $Df(x)$, so the orthogonal projection $P^{\perp}(x, \dot{x})$ to the orthogonal complement of $\ker Df(x)$ has the formula[1] $\dot{x} \mapsto Df(x)^t\big(Df(x)Df(x)^t\big)^{-1}Df(x)\dot{x}$. P is C^∞ since it has the formula $P(x, \dot{x}) = \dot{x} - P^{\perp}(x, \dot{x})$.

Since X is C^∞, by definition it extends to a C^∞ vector field $\tilde{X}$ defined on an open subset $U \subseteq \mathbb{R}^m$ containing M. For all $x \in M$, $X(x) \in \ker Df(x)$, since X is a vector field on M, so $P\big(x, X(x)\big) = X(x)$ and $x \mapsto P\big(\tilde{X}(x)\big)$ is also an extension of X. So we may assume that X has an extension $\tilde{X}$ to an open subset U of M such that $\tilde{X}(x) \in \ker Df(x)$ for all $x \in U$.

Let $\tilde{x}(t)$ be the maximal integral curve of $\tilde{X}$ starting at x_0 at time t_0. Then f is constant on $\tilde{x}(t)$ since

$$\frac{d}{dt} f\big(\tilde{x}(t)\big) = Df\big(x(t)\big)\frac{d\tilde{x}}{dt} = Df\big(x(t)\big)\tilde{X}(t) = 0.$$

Thus $\tilde{x}(t) \in M = f^{-1}(a)$ for all t, and, because $\tilde{X}$ extends X, $\tilde{x}(t)$ is an integral curve of X. Any integral curve of X staring at x_0 at time t_0 is also an integral curve of $\tilde{X}$ and hence is a restriction of $\tilde{x}$. Thus $\tilde{x}$ is a maximal integral curve of X. This is the unique maximal integral curve of X starting at x_0 at time t_0 because any two such curves would be restrictions of each other. ∎

Remark 6 The proof shows that if X is a vector field defined on U, where U is the domain of the submersion defining M, and if $X(x) \in TM$ for all $x \in M$, then the integral curves of X starting in M remain in M over the whole interval of their definition.

1.4 Lie groups and actions

In the sense of this thesis, a *Lie group* is a manifold G which is also a subgroup of the $p \times p$ invertible real matrices $GL(p)$. If G is a Lie group, then the vector space

[1] There is a unique $b \in \mathbb{R}^n$ such that $P^{\perp}(x, \dot{x}) = Df(x)^t b$ since the columns of $Df(x)^t$ are a basis of $\ker Df(x)^{\perp}$. The condition that $P^{\perp}(x, \dot{x})$ is the orthogonal projection to $\ker Df(x)^{\perp}$ is that $\dot{x} - P^{\perp}(x, \dot{x}) \in \ker Df(x)$ i.e.

$$Df(x)\dot{x} - Df(x)P^{\perp}(x, \dot{x}) = Df(x)\dot{x} - Df(x)Df(x)^t b = 0. \tag{1.1}$$

$Df(x)Df(x)^t$ is invertible because it has trivial kernel: if $Df(x)Df(x)^t y = 0$ then $Df(x)^t y \in \ker Df(x) \cap \ker Df(x)^{\perp}$ so $Df(x)^t y = 0$ and then $y = 0$ since $Df(x)^t$ is injective. Solving (1.1) for b gives $b = \big(Df(x)Df(x)^t\big)^{-1}Df(x)\dot{x}$ and hence $P^{\perp}(x, \dot{x}) = Df(x)^t\big(Df(x)Df(x)^t\big)^{-1}\dot{x}$.

of $p \times p$ matrices defined by

$$\mathfrak{g} = \bigl\{\xi \bigm| (\mathbf{1}, \xi) \in T_{\mathbf{1}}G\bigr\}$$

is called the *Lie algebra* of G.

Since any open subset of the $p \times p$ matrices is a manifold in the sense of this thesis, $GL(p)$ itself is a Lie group in the sense of this thesis. Its Lie algebra is the set of $p \times p$ matrices.

The *matrix exponential* of $\xi \in \mathfrak{g}$ is defined by the (absolutely convergent) series

$$\exp \xi = \mathbf{1} + \frac{1}{1!}\xi + \frac{1}{2!}\xi^2 + \frac{1}{3!}\xi^3 + \dots.$$

The curve $g(t) = \exp(\xi t)$ satisfies the differential equation

$$\frac{dg}{dt} = g\xi. \tag{1.2}$$

For all g, $(g, g\xi) \in T_gG$. This is implied by Remark 3: $\xi = h'(0)$ for some curve $h(t) \in G$, $gh(t) \in G$ since G is closed under product, and $g\xi = \bigl(gh(t)\bigr)'(0)$. So, the vector field corresponding to the differential equation (1.2) is everywhere tangent to G, and since the curve $\exp(\xi t)$ satisfies (1.2), Remark (6) implies $\exp(\xi t) \in G$ for all $t \in \mathbb{R}$. Consequently, the exponential mapping takes values in G for all $\xi \in \mathfrak{g}$.

The ordinary real number exponential e^x has the same series expansion as $\exp \xi$. The common identities $e^x e^{-x} = 1$, and $e^{x+y} = e^x e^y$, written out as series, imply the validity of the corresponding series manipulations that verify the identities

$$\exp\bigl(\xi(t+s)\bigr) = \exp(\xi t)\exp(\xi s), \qquad \xi \in \mathfrak{g},\ s, t \in \mathbb{R},$$

and

$$\exp(\xi)\exp(-\xi) = \mathbf{1}, \qquad \xi \in \mathfrak{g}.$$

In particular, $\bigl(\exp(\xi)\bigr)^{-1} = \exp(-\xi)$. However, it should be noted that, as a general identity, $\exp(\xi+\eta) = (\exp \xi)(\exp \eta)$ is false due to the general lack of commutativity of the matrices ξ and η. In fact, the identity fails in $GL(2)$ with

$$\xi = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \eta = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

as is easily verified.

A *linear representation* of G on $\mathbb{R}^m$ is a map $G \times \mathbb{R}^m \to \mathbb{R}^m$ denoted $(g, x) \mapsto gx$, such that

1. $\mathbf{1}x = x$ for all $x \in \mathbb{R}^m$;

2. $g(hx) = (gh)x$ for all $g, h \in G$ and all $x \in \mathbb{R}^m$;

3. for all $g \in G$, the mapping $x \mapsto gx$ is a linear map of $\mathbb{R}^m$.

Equivalently, we have a matrix $\Psi_g \in GL(m)$ for each $g \in G$ and $g \mapsto \Psi_g$ is a group homomorphism.

Define the *adjoint* Ad_g of $g \in G$ by

$$\operatorname{Ad}_g \xi = g\xi g^{-1}, \qquad g \in G,\ \xi \in \mathfrak{g}.$$

Directly from the series definition of $\exp(\xi t)$,

$$g \exp(\xi t) g^{-1} = \exp(\operatorname{Ad}_g \xi t),$$

so $\exp(\operatorname{Ad}_g \xi t) \in G$ for all t. Differentiating at $t = 0$ gives $\operatorname{Ad}_g \xi \in \mathfrak{g}$ by Remark 3. So Ad is a linear representation, called the *adjoint representation*, of G on $\mathfrak{g}$ since

1. $\operatorname{Ad}_{\mathbf{1}} \xi = \mathbf{1}\xi\mathbf{1}^{-1} = \xi$;

2. $\operatorname{Ad}_g \operatorname{Ad}_h \xi = g(h\xi h^{-1})g^{-1} = (gh)\xi(gh)^{-1} = \operatorname{Ad}_{gh} \xi$;

3. $\operatorname{Ad}_g(s\xi + t\eta) = g(s\xi + t\eta)g^{-1} = sg\xi g^{-1} + tg\xi g^{-1} = s \operatorname{Ad}_g \xi + t \operatorname{Ad}_g \eta$.

Define the *coadjoint representation* of G on the dual space $\mathfrak{g}^*$ by

$$\langle \operatorname{CoAd}_g \mu, \xi \rangle = \langle \mu, \operatorname{Ad}_{g^{-1}} \xi \rangle, \qquad g \in G, \mu \in \mathfrak{g}^*$$

where the evaluation of dual vectors $\mu \in \mathfrak{g}^*$ on vectors $\xi \in \mathfrak{g}$ is denoted by $\langle \mu, \xi \rangle$. This is a linear representation of G on the dual space $\mathfrak{g}^*$.

If G is commutative, then

$$\operatorname{Ad}_g \xi = \left.\frac{d}{dt}\right|_{t=0} g^{-1} \exp(\xi t) g = \left.\frac{d}{dt}\right|_{t=0} g^{-1} g \exp(\xi t) = \xi.$$

Thus the adjoint and coadjoint actions are trivial for commutative groups.

If $\xi, \eta \in \mathfrak{g}$ then $\operatorname{Ad}_{\exp \xi t} \eta \in \mathfrak{g}$. Differentiating at $t = 0$ gives

$$\left.\frac{d}{dt}\right|_{t=0} \exp(\xi t)\eta \exp(-\xi t) = \xi\eta - \eta\xi \in \mathfrak{g}.$$

This shows that $\mathfrak{g}$ is closed under the *Lie bracket* operation $[\xi, \eta] = \xi\eta - \eta\xi$. As is easily verified, the Lie bracket is bilinear and anticommutative, and satisfies the *Jacobi identity*

$$[\xi, [\eta, \nu]] + [\eta, [\nu, \xi]] + [\nu, [\xi, \eta]] = 0, \qquad \xi, \eta, \nu \in \mathfrak{g}.$$

Generally, a vector space endowed with such a bracket is called a *Lie algebra*. Thus, the term "Lie algebra of G" is an appropriate one for the vector space $\mathfrak{g}$. If G is commutative then its Lie bracket is zero (i.e. $[\xi, \eta] = 0$ for all $\xi, \eta \in \mathfrak{g}$) since its adjoint representation is trivial and therefore differentiates to zero.

If $M \subseteq \mathbb{R}^m$ is a manifold, then by an *action of G on M* we will mean a linear representation of G on $\mathbb{R}^m$ such that $gx \in M$ for all $g \in G$ and $x \in M$. The *orbit* of $x \in M$ is the set $Gx = \{gx \mid g \in G\}$. The *isotropy subgroup* of $x \in M$ is the subgroup $G_x = \{g \mid gx = x\}$. The group action is *free* if the isotropy group of every $x \in M$ is trivial.

For any $x \in \mathbb{R}^m$ and any $\xi \in \mathfrak{g}$, define the *infinitesimal generator of ξ at x* by

$$I(\xi, x) = \left.\frac{d}{dt}\right|_{t=0} \exp(\xi t)x. \tag{1.3}$$

By Remark 3, $\big(x, I(\xi x)\big) \in T_xM$ for all $x \in M$. I is a bilinear map from $\mathfrak{g} \times \mathbb{R}^m$ to $\mathbb{R}^m$. We will use the notation I_ξ for the linear operator on $\mathbb{R}^m$ obtained from I by fixing some ξ. Equation (1.3) explicitly writes $I_\xi x$ as the derivative of a curve in M, so I_ξ is a vector field on M, and, again from equation (1.3), the integral curve of this vector field starting at $x \in M$ at time $t = 0$ is $\exp(\xi t)x$. The infinitesimal generator satisfies the equation

$$I_\xi(gx) = \left.\frac{d}{dt}\right|_{t=0} \exp(\xi t)(gx) = \left.\frac{d}{dt}\right|_{t=0} g\big(g^{-1}\exp(\xi t)g\big)x = gI_{\mathrm{Ad}_{g^{-1}}\xi}(x). \tag{1.4}$$

1.5 Example: S^2 and $SO(2)$

As an example, we consider $f\colon \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}$ by $f(x) = |x|^2$. The derivative of this map is the 1×3 matrix $Df(x) = 2x$, which has full rank if $x \neq 0$. So f is a submersion since its domain does not include $0 \in \mathbb{R}^3$, and we can consider the manifold $S^2 = f^{-1}(1)$ i.e. the 2-sphere in $\mathbb{R}^3$.

$Tf(x,\dot{x}) = (|x|^2, 2x\cdot\dot{x})$ is a submersion and the tangent bundle to S^2 is the set $(x,\dot{x}) \in (\mathbb{R}^3)^2$ such that $Tf(x,\dot{x}) = (|x|^2, x\cdot\dot{x}) = (1,0)$ i.e.

$$TS^2 = \big\{(x,\dot{x}) \;\big|\; |x| = 1, x\cdot\dot{x} = 0\big\},$$

which is the set of vectors with base points in S^2 which are tangent to S^2 at the base points.

Consider the function h, defined on the 2×2 matrices with positive determinant, by $h(A) = A^tA$. We view h as taking values in the (3 dimensional) vector space of 2×2 *symmetric* matrices. The derivative of h is

$$Dh(A)\dot{A} = \frac{d}{dt}\bigg|_{t=0}(A+t\dot{A})^t(A+t\dot{A}) = A^t\dot{A} + \dot{A}^tA = A^t\dot{A} + (A^t\dot{A})^t. \tag{1.5}$$

$\ker Dh(A)$ is the set of $\dot{A}$ such that $A^t\dot{A}$ is antisymmetric, or equivalently, those $\dot{A}$ for which $\dot{A} = (A^t)^{-1}X$, where X is antisymmetric. Thus $\ker Dh(A)$ has dimension 1, so $Dh(A)$ is onto since image $Dh(A)$ has dimension $4-1=3$, the same dimension as the codomain of h. Thus, h is a submersion. $h^{-1}(\mathbf{1})$ is easily verified to be a subgroup of $GL(2)$ so $h^{-1}(\mathbf{1})$ is a Lie group in the sense of this thesis.

The 2×2 matrix

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

satisfies $h(A) = \mathbf{1}$ if and only if $ad-bc > 0$, $a^2+b^2 = 1$, $c^2+d^2 = 1$, and $ab+cd = 0$. Thus (a,c) and (b,d) are orthogonal unit vectors in $\mathbb{R}^2$, and there is some θ such that

$$(a,c) = (\cos\theta, \sin\theta), \quad (b,d) = \pm(-\sin\theta, \cos\theta).$$

$ad-bc > 0$ implies the positive sign, so

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \tag{1.6}$$

and one sees that $h^{-1}(\mathbf{1})$ is the group of proper rotations of the plane. This group is denoted $SO(2)$.

From (1.5), the derivative of h at $\mathbf{1}$ is $\dot{A} \mapsto \dot{A} + \dot{A}^t$. The Lie algebra$so(2)$ of $SO(2)$ is the kernel of this and thus consists of the antisymmetric 2×2 matrices.

This one dimensional vector space may be identified with $\mathbb{R}$ by the linear bijection $\xi \leftrightarrow \mathbf{J}\xi$ where $\mathbf{J}$ is the matrix

$$\mathbf{J} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{1.7}$$

Since $\mathbf{J}^2 = -\mathbf{1}$, the exponential of $\mathbf{J}\xi$ this matrix is easily calculated:

$$\begin{aligned} \exp \mathbf{J}\xi &= \mathbf{1} + \frac{\xi}{1!}\mathbf{J} - \frac{\xi^2}{2!}\mathbf{1} - \frac{\xi^3}{3!}\mathbf{J} + \frac{\xi^4}{4!}\mathbf{1} + \cdots \\ &= \begin{bmatrix} \cos\xi & -\sin\xi \\ \sin\xi & \cos\xi \end{bmatrix}. \end{aligned}$$

In accord with the general theory, $\exp\xi \in SO(2)$ for all $\xi \in so(2)$. The dual of $so(2)$ can also be identified with $\mathbb{R}$ with the understanding that $\mu \in \mathbb{R}$ pairs with $\xi \in so(2) = \mathbb{R}$ by ordinary multiplication of numbers: $\langle \mu, \xi \rangle = \mu\xi$. Since $SO(2)$ is commutative, its adjoint and coadjoint representations are trivial, and its Lie bracket is zero.

For $x \in \mathbb{R}^3$, the multiplication rule

$$Ax = \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix} x$$

is a linear representation of $SO(2)$ on $\mathbb{R}^3$. This representation of $SO(2)$ corresponds to the proper rotations of $\mathbb{R}^3$ about the axis $\mathbf{k} = (0, 0, 1)$, and it gives an action of $SO(2)$ of S^2. The curves $\exp(\xi t)x$ are rotations of angular speed ξ around the axis $\mathbb{R}\mathbf{k}$. The infinitesimal generator of $\xi \in so(2)$ is

$$I_\xi x = \left.\frac{d}{dt}\right|_{t=0} \exp(\xi t)x = \begin{bmatrix} \xi\mathbf{J} & 0 \\ 0 & 0 \end{bmatrix} x = \xi\mathbf{k} \times x. \tag{1.8}$$

The orbits of the action are the intersections of horizontal planes with S^2. The action is free except at $x = \pm\mathbf{k}$, where the isotropy group is $SO(2)$ itself.

Chapter 2

Lagrangian Systems

We will consider certain mechanical systems with mathematical models specified by a manifold $Q \subseteq \mathbb{R}^m$, called the *configuration space*, and a C^∞ function

$$L\colon \big\{(q,v)\big\} = \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R},$$

called the *Lagrangian*. We will assume $Q = f^{-1}(0)$ where $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is a given submersion. It is understood that the physical configurations of the mechanical system are, to an adequate approximation, in bijective correspondence with the points of Q. The manifold $P = TQ$ is called *phase space*. The elements of P are called the *states* of the system. f is often referred to as a *constraint*.

2.1 Variational Principle

The *action* of a C^1 curve $q(t) \in Q$, $t \in [a,b]$, is by definition

$$S\big(q(t)\big) = \int_a^b L \circ \big(q(t), q'(t)\big)\, dt.$$

A *variation* of the curve q is a family of curves $q_\epsilon(t) \in Q$, $t \in [a,b]$, such that $(t,\epsilon) \mapsto q_\epsilon(t)$ is C^1 and $q_0(t) = q(t)$. Given such a variation, we define

$$\delta q(t) = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} q_\epsilon(t), \qquad dS\big(q(t)\big)\cdot \delta q(t) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} S\big(q_\epsilon(t)\big).$$

The action is *critical at* $q(t)$ if

$$dS\big(q(t)\big)\cdot \delta q(t) = 0,$$

for all variations q_ϵ which have values in Q (i.e. $q_\epsilon(t) \in Q$ for all t, ϵ) and are such that $q_\epsilon(a) = q(a)$ and $q_\epsilon(b) = q(b)$. The physical evolutions of the system are modeled by the curves $q(t)$ for which the action is critical, given fixed endpoints. The identification of the physical evolutions with the critical curves is called *Hamilton's principle*.

An integration by parts gives the expression

$$
\begin{aligned}
dS\big(q(t)\big) \cdot \delta q(t) &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \int_a^b L\left(q_\epsilon^i(t), \frac{dq_\epsilon^i}{dt}(t)\right) dt \\
&= \int_a^b \left(\frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial v^i}\, \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \frac{dq_\epsilon^i}{dt}(t)\right) dt \\
&= \int_a^b \left(\frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial v^i}\frac{d}{dt}\, \frac{dq_\epsilon^i(t)}{d\epsilon}\bigg|_{\epsilon=0}\right) dt \\
&= \int_a^b \left(\frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial v^i}\frac{d(\delta q^i)}{dt}\right) dt \\
&= \frac{\partial L}{\partial v^i}\delta q^i\bigg|_a^b + \int_a^b \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial v^i}\right) \delta q^i \, dt.
\end{aligned} \tag{2.1}
$$

The first term vanishes if $\delta q(t)$ satisfies the conditions $\delta q(a) = \delta q(b) = 0$. Also, if $f\big(q_\epsilon(t)\big) = 0$ then

$$
0 = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} f\big(q_\epsilon(t)\big) = Df\big(q(t)\big)\delta q(t),
$$

i.e. $\delta q(t) \in T_{q(t)}Q$ for all t. Thus, $q(t)$ is an evolution if, for all t,

$$
\left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial v^i}\right)\big(q(t), q'(t)\big)v^i = 0 \tag{2.2}
$$

for all $v \in \mathbb{R}^m$ such that $Df\big(q(t)\big)v = 0$. Conversely, $q(t)$ is not an evolution if (2.2) fails for some t and some such v. Indeed, if (2.2) fails for t and v such that $Df\big(q(t)\big)v = 0$, then using Theorem 1, we can realize this v as $\delta q(t)$ from a variation $q_\epsilon(t) \in Q$ such that the left side of (2.2) is either positive near t and zero otherwise, or negative near t and zero otherwise, and the integral in (2.1) will not vanish for this variation.

For any $q \in Q$ and for each $\lambda \in \mathbb{R}^n$, consider the linear map

$$
\lambda \mapsto \lambda^t Df(q). \tag{2.3}
$$

This map is a bijection to the subspace $\big(\ker Df(q)\big)^\perp$ because the rows of $Df(q)$

are a basis of $\big(\ker Df(q)\big)^{\perp}$. Since condition (2.2) is exactly

$$\left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial v^i}\right)\big(q(t), q'(t)\big) \in \Big(\ker Df\big(q(t)\big)\Big)^{\perp} \tag{2.4}$$

it follows that $q(t)$ is critical if and only if for each t there is a (Lagrange multiplier) $\lambda \in \mathbb{R}^n$ such that

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial v^i} = \lambda_a \frac{\partial f^a}{\partial q^i}, \qquad v^i = \frac{dq^i}{dt}, \qquad f^a\big(q(t)\big) = 0, \tag{2.5}$$

where the components of f are denoted f^a, $1 \le a \le n$.

The *energy* of the Lagrangian system is defined to be the function

$$E = \frac{\partial L}{\partial v^i} v^i - L.$$

If $q(t) \in Q$ is an evolution then

$$\begin{aligned}
\frac{dE}{dt} &= \frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right) v^i + \frac{\partial L}{\partial v^i}\frac{dv^i}{dt} - \frac{\partial L}{\partial q^i}\frac{dq^i}{dt} - \frac{\partial L}{\partial v^i}\frac{dv^i}{dt} && \text{(Chain rule)}\\
&= \left(\frac{\partial L}{\partial q^i} - \lambda_a \frac{\partial f^a}{\partial q^i}\right)\frac{dq^i}{dt} - \frac{\partial L}{\partial q^i}\frac{dq^i}{dt} && \text{(equation (2.5))}\\
&= -\lambda_a \frac{df^a}{dt} && \text{(Chain rule)}\\
&= 0. && f\big(q(t)\big) = 0
\end{aligned}$$

Thus, energy is constant along the evolutions.

Define the quantity

$$F^2 L(q,v)(w,w') = D^2 L(q,v)\big((0,w),(0,w')\big), \quad (q,w),(q,w') \in T_q Q.$$

For each $(q,v) \in TQ$, $F^2 L(q,v)(w,w')$ is the bilinear form obtained from L by differentiating L in the variables v^i. L is called *regular* at $(q,v) \in T_q Q$ if, for all $(q,w) \in T_q Q$, $F^2 L(q,v)(w,w') = 0$ for all $(q,w') \in T_q Q$ implies $w = 0$. Equivalently, L is regular if, for each $q \in Q$, the matrix of the bilinear form $F^2 L(q,v)$ (obtained from any basis of $\ker Dh(q)$) is nonsingular. Given that L is regular, Equations (2.5) can be arranged so $q(t)$ is an evolution if and only if $\big(q(t), q'(t)\big)$ is an integral curve of a vector field on $TTQ \subseteq (R^{2m})^2$. To see this, expand the time derivative in the first equation of (2.5) and differentiate twice the third equation, to obtain

$$\frac{\partial^2 L}{\partial v^i \partial v^j}\frac{dv^j}{dt} + \frac{\partial f^a}{\partial q^i}\lambda_a = \frac{\partial L}{\partial q^i} - \frac{\partial^2 L}{\partial v^i \partial q^j} v^j, \tag{2.6}$$

$$\frac{\partial f^a}{\partial q^i}\frac{dv^i}{dt} = -\frac{\partial^2 f^a}{\partial q^i \partial q^j} v^i v^j. \tag{2.7}$$

This is a system of linear equations for dv/dt and λ, with coefficient matrix

$$\begin{bmatrix} D^2 L(q,v) & Df(q)^t \\ Df(q) & 0 \end{bmatrix}. \tag{2.8}$$

This matrix is nonsingular if L is regular. Indeed, $w, \delta\lambda$ is in the kernel of (2.8) if and only if

$$D^2 L(q,v)w + Df(q)^t \delta\lambda = 0, \qquad Df(q)w = 0.$$

For any $w' \in \ker Df(q)$,

$$(w')^t \big(D^2 L(q,v)\big) w = D^2 L(q,v)(w,w') = -w' \cdot Df(q)^t \delta\lambda = -Df(q)w' \cdot \delta\lambda = 0,$$

so $w = 0$. So $Df(q)^t \delta\lambda = 0$ so $\delta\lambda = 0$ since $Df(q)^t$ is injective (because $Df(q)$ is surjective). Thus equations (2.6) and (2.7) can be solved uniquely and smoothly for dv^i/dt and λ_a, and can be written

$$\frac{dv^i}{dt} = X^i(q,v), \quad \frac{dq^i}{dt} = v^i.$$

The vector field $(q,v) \mapsto \big(v, X(q,v)\big)$ is tangent to TTQ because it satisfies (2.7). Therefore, by Theorem 5, for any given $(q_0, v_0) \in TQ$, and any t_0 there is a unique maximal evolution $q(t) \in Q$ such that $q(t_0) = q(t)$ and $q'(t_0) = v_0$.

2.2 Symmetry

General Lagrangian systems model diverse and complicated physical phenomena, including the classical dynamics of molecules and the motion of the planets. We are concerned not with generic Lagrangian systems, but rather with systems which admit a continuous symmetry, and for which the Lagrangian is the difference between a kinetic and potential energy.

There will be a Lie group G and a linear representation of G of $\mathbb{R}^m$ such that $gq \in Q$ for all $g \in G$ and $q \in Q$. The submersion $f\colon \mathbb{R}^m \to \mathbb{R}^n$ which defines the configuration space $Q = f^{-1}(0)$ will be G invariant: $f(gq) = f(q)$ for all $q \in Q$ and all $g \in G$. Thus, there will be an action of G on Q in the sense of Section 1.4. For each $q \in \mathbb{R}^m$ there will be a positive definite symmetric matrix $M(q)$ which defines metrics $\langle v, w\rangle_q = v^t M w$ on $\mathbb{R}^m$, called the *kinetic energy metric*. The metric will be invariant under the representation, so that

$$\langle gv, gw\rangle_{gq} = \langle v, w\rangle_q, \qquad v, w \in \mathbb{R}^m,\ g \in G.$$

The Lagrangian L will be given by

$$L(q, v) = \frac{1}{2}|v|_q^2 - V(q) = \frac{1}{2}v^t M(q) v - V(q), \tag{2.9}$$

where V is a C^∞ function, called the *potential energy* on $\mathbb{R}^m$ which is invariant under the action of G i.e. $V(gq) = V(q)$ for all $q \in Q$, $g \in G$.

For the Lagrangians of the form (2.9) i.e.

$$L = \frac{1}{2}M_{kl}(q)v^k v^l - V(q),$$

and if M is constant, equations (2.5) simplify:

$$\begin{aligned}
\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial v^i} &= -\frac{\partial V}{\partial q^i} - \frac{d}{dt}\left(\frac{1}{2}M_{kl}\delta_i^k v^j + \frac{1}{2}M_{kl}v^k\delta_i^l\right) \\
&= -\frac{\partial V}{\partial q^i} - \frac{d}{dt}(M_{ij}v^j).
\end{aligned}$$

Thus $\big(q(t), v(t)\big)$ is an evolution if and only if

$$\frac{d}{dt}(Mv) = -\nabla(V + \lambda\cdot f), \qquad \frac{dq}{dt} = v, \qquad f\big(q(t)\big) = 0, \tag{2.10}$$

which are Newton equations of elementary mechanics.

For such a Lagrangian as (2.9), the energy is

$$\begin{aligned}
E(q, v) &= \frac{1}{2}\frac{\partial}{\partial v^j}(M_{kl}v^k v^l)v^j - L \\
&= \frac{1}{2}M_{kl}\delta_j^k v^l v^j + \frac{1}{2}M_{kl}v^k\delta_j^l v^j - \frac{1}{2}M_{kl}v^k v^l + V(q) \\
&= \frac{1}{2}M_{kl}v^k v^l + V(q) \\
&= \frac{1}{2}|v|^2 + V(q)
\end{aligned}$$

i.e. the sum of the kinetic and potential energies.

Under the conditions assumed, $L(gq, gv) = L(q, v)$ i.e. the Lagrangian L is invariant, since

$$L(gq, gv) = \frac{1}{2}\langle gv, gv\rangle - V(gq) = \frac{1}{2}\langle v, v\rangle - V(q) = L(q, v).$$

The action S is also invariant, meaning that for any $g \in G$ and any curve $q(t)$, $S\big(gq(t)\big) = S\big(q(t)\big)$. This follows since

$$\begin{aligned} S\big(gq(t)\big) &= \int_a^b L \circ \big(gq(t), \big(gq(t)\big)'\big)\, dt \\ &= \int_a^b L \circ \big(gq(t), gq'(t)\big)\, dt \\ &= \int_a^b L \circ \big(q(t), q'(t)\big)\, dt. \end{aligned}$$

If $q(t)$ is an evolution and $g \in G$, then $gq(t)$ is also an evolution: given any variation $q_\epsilon(t)$ of $gq(t)$,

$$\begin{aligned} \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S\big(q_\epsilon\big) &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_a^b L\big(q_\epsilon(t), q_\epsilon'(t)\big) \\ &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_a^b L\big(g^{-1}q_\epsilon(t), g^{-1}q_\epsilon'(t)\big) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S\big(g^{-1}q_\epsilon\big) = 0 \end{aligned}$$

since $g^{-1}q_\epsilon(t)$ is a variation of $q(t)$, which is an evolution.

2.3 Relative equilibria

We seek pairs $(q_e, \xi_e) \in Q \times \mathfrak{g}$ such that the curve $q(t) = \exp(\xi_e t)q_e$ is an evolution. These points of $Q \times \mathfrak{g}$ are called *relative equilibria*. They correspond to special, symmetry related evolutions of the system.

A convenient necessary and sufficient condition for (q_e, ξ_e) to be a relative equilibrium can be derived directly from the variational principle stated in Section 2.1. Define the function

$$V_\xi(q) = V(q) - \frac{1}{2}|I_\xi q|^2,$$

which is the negative of the Lagrangian evaluated on the infinitesimal generator. This function is called the *amended potential*. Variations $q_\epsilon(t)$ of $q(t) = \exp(\xi_e t)q_e$ and the maps $\tilde{q}_\epsilon(t)$ such that $\tilde{q}_0(t) = q_e$ are in bijective correspondence by

$$\tilde{q}_\epsilon(t) = \exp(-\xi_e t)q_\epsilon(t), \qquad q_\epsilon(t) = \exp(\xi_e t)\tilde{q}_\epsilon(t).$$

so the variations $q_\epsilon(t)$ in the variational principle may be assumed to be of the form $\exp(\xi_e t)\tilde{q}_\epsilon(t)$ such that

$$\tilde{q}_0(t) = q_e, \quad \tilde{q}_\epsilon(a) = q_e, \quad \tilde{q}_\epsilon(b) = q_e.$$

The invariance of the Lagrangian can be used to calculate the derivative of the action with respect to such variations:

$$
\begin{aligned}
dS\big(q(t)\big)\delta q &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \int_a^b L\left(\exp(\xi_e t)\tilde{q}_\epsilon(t), \frac{d}{dt}\exp(\xi_e t)\tilde{q}_\epsilon(t)\right)\, dt \\
&= \int_a^b \frac{d}{d\epsilon}\bigg|_{\epsilon=0} L\left(\exp(\xi_e t)\tilde{q}_\epsilon(t), \exp(\xi_e t) I_{\xi_e}\tilde{q}_\epsilon(t) + \exp(\xi_e t)\frac{d}{dt}\tilde{q}_\epsilon(t)\right)\, dt \\
&= \int_a^b \frac{d}{d\epsilon}\bigg|_{\epsilon=0} L\left(\tilde{q}_\epsilon(t), I_{\xi_e}\tilde{q}_\epsilon(t) + \frac{d}{dt}\tilde{q}_\epsilon(t)\right)\, dt \\
&= \int_a^b \left(\frac{d}{d\epsilon}\bigg|_{\epsilon=0} L\left(\tilde{q}_0(t), I_{\xi_e}\tilde{q}_0(t) + \frac{d}{dt}\tilde{q}_\epsilon(t)\right)\right. \\
&\qquad\qquad \left. + \frac{d}{d\epsilon}\bigg|_{\epsilon=0} L\left(\tilde{q}_\epsilon(t), I_{\xi_e}\tilde{q}_\epsilon(t) + \frac{d}{dt}\tilde{q}_0(t)\right)\right)\, dt \\
&= \int_a^b \left(\frac{\partial L}{\partial v^i}(q_e, I_{\xi_e}q_e)\ \frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\frac{\partial}{\partial t}\tilde{q}^i_\epsilon(t) + \frac{d}{d\epsilon}\bigg|_{\epsilon=0} L\big(\tilde{q}_\epsilon(t), I_{\xi_e}\tilde{q}_\epsilon(t)\big)\right)\, dt \\
&= \int_a^b \left(\frac{d}{dt}\frac{\partial L}{\partial v^i}(q_e, I_{\xi_e}q_e)\ \frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\tilde{q}^i_\epsilon(t) - \frac{d}{d\epsilon}\bigg|_{\epsilon=0} V_{\xi_e}\circ\tilde{q}_\epsilon(t)\right)\, dt \\
&= \left(\frac{\partial L}{\partial v^i}(q_e, I_{\xi_e}q_e)\ \frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\tilde{q}^i_\epsilon(t)\right)\bigg|_a^b - \int_a^b \left(\frac{d}{d\epsilon}\bigg|_{\epsilon=0} V_{\xi_e}\circ\tilde{q}_\epsilon(t)\right)\, dt \\
&= -\int_a^b \frac{\partial V_{\xi_e}}{\partial q^i}(q_e)\delta\tilde{q}^i(t)\, dt
\end{aligned}
$$

where

$$\delta\tilde{q}(t) = \frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\tilde{q}_\epsilon(t)$$

satisfies $\delta\tilde{q}(a) = \delta\tilde{q}(b) = 0$ and $Df(q_e)\delta q = 0$, but is otherwise arbitrary. Thus (q_e, ξ_e) is a relative equilibrium if and only if

$$q_e \text{ is a critical point of } V_{\xi_e}|f^{-1}(0). \tag{2.11}$$

Without Lagrange multipliers, (2.11) is

$$\nabla\big(V_{\xi_e}(q_e)\big)\cdot\delta q = 0 \quad \text{for all} \quad \delta q \in \ker Df(q_e), \tag{2.12}$$

or equivalently

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} V_{\xi_e}(q_e + \epsilon\delta q) = 0 \quad \text{for all} \quad \delta q \in \ker Df(q_e). \tag{2.13}$$

With Lagrange multipliers, (2.11) is

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} (V_{\xi_e} + \lambda\cdot f)(q_e + \epsilon\delta q) = 0 \quad \text{for all} \quad \delta q \in \mathbb{R}^m \tag{2.14}$$

for some $\lambda \in \mathbb{R}^n$. In the case where the submersion f which defines $Q = f^{-1}(0)$ is of the form $f = (f_1, f_2)\colon R^m \to \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, we can explicitly hold the tangent vectors in $f_1^{-1}(0)$ and use Lagrange multipliers for the f_2 part, and (2.11) becomes

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} (V_{\xi_e} + \lambda \cdot f_2)(q_e + \epsilon\delta q) = 0 \quad \text{for all} \quad \delta q \in \ker Df_1(q_e) \tag{2.15}$$

for some $\lambda \in \mathbb{R}^{n_2}$.

If (q_e, ξ_e) is a relative equilibrium, then $\exp(\xi_e t)q_e$ is an evolution, $g\big(\exp(\xi_e t)q_e\big)$ is an evolution, and

$$g\big(\exp(\xi_e t)q_e\big) = \big(g \exp(\xi_e t)g^{-1}\big)(gq_e) = \exp(g\xi_e g^{-1}t)(gq_e) = \exp(\operatorname{Ad}_g \xi t)(gq_e).$$

So, if (q_e, ξ_e) is a relative equilibrium and $g \in G$, then $(gq_e, \operatorname{Ad}_g \xi_e)$ is also a relative equilibrium. Let $S \subset Q$ be such that S intersects each group orbit Gq, $q \in Q$, exactly once. Since gq_e is a relative equilibrium when q_e is, without loss of generality, one can when finding relative equilibria, restrict the search to S. All relative equilibria can then be obtained by multiplication of the relative equilibria in S by elements in G.

2.4 Momentum

As we have already noted, for the systems we are considering, the action is invariant: given any $g \in G$,

$$S\big(gq(t)\big) = \int_a^b L \circ \big(gq(t), (gq)'(t)\big)\, dt = \int_a^b L \circ \big(q(t), q'(t)\big)\, dt = S\big(q(t)\big).$$

In particular, if $\xi \in \mathfrak{g}$ and $g = \exp(\xi\epsilon)$, then

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S\big(\exp(\xi\epsilon)q(t)\big) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S\big(q(t)\big) = 0. \tag{2.16}$$

Suppose that $q(t)$ is an evolution. Then $q_\epsilon(t) = \exp(\xi\epsilon)q(t)$ is a variation

$$\delta q = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \exp(\xi\epsilon)q(t) = I_\xi q(t).$$

From (2.1) and (2.16),

$$0 = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S\big(\exp(\xi\epsilon)q(t)\big) = dS\big(q(t)\big) \cdot \delta q(t) = \left.\frac{\partial L}{\partial v^i}\delta q^i\right|_a^b = \left. v^t M(q) I_\xi q(t)\right|_a^b.$$

If we define the function $J_\xi \colon TQ \to \mathbb{R}$ by

$$J_\xi(q, v) = v^t M(q) I_\xi q = \langle v, I_\xi q \rangle_q, \tag{2.17}$$

then, for any a, b, J_ξ has the same value on an evolution $q(t)$ at $t = a$ as it has at $t = b$. J_ξ is called the *momentum generated by* ξ. Because $\xi \in \mathfrak{g}$ is arbitrary, and the dependence of J_ξ on ξ is linear, there are as many independent momenta as there are dimensions of symmetry. The mapping $J \colon TQ \to \mathfrak{g}^*$ defined by

$$J(q, v) \cdot \xi = J_\xi(q, v) \tag{2.18}$$

is called the *momentum mapping* or just the *momentum* for the system. The momentum of (q, v) is by definition an element of the dual of the Lie algebra of the symmetry group of the system.

Suppose a system is in state $(q, v) \in TQ$. That state has momentum $J(q, v) \in \mathfrak{g}^*$. If the system is re-oriented to state (gq, gv), where $g \in G$, then the momentum should change in a way that depends on g. The momentum map $J \colon TQ \to \mathfrak{g}^*$ satisfies an important equivariance property which captures this expectation. Because of equation (1.4), the momentum of the state (gq, gv) satisfies

$$J_\xi(gq, gv) = \langle gv, I_\xi(gq) \rangle_{gq} = \langle gv, g I_{\mathrm{Ad}_{g^{-1}} \xi} \rangle_{gq} = \langle v, I_{\mathrm{Ad}_{g^{-1}}\xi}(q) \rangle_q = J_{\mathrm{Ad}_{g^{-1}}\xi}(q, v),$$

so, since ξ is arbitrary,

$$J(gq, gv) = \mathrm{CoAd}_g \, J(q, v).$$

2.5 Particle on a surface of revolution

As an example, consider a flat graph $z = F(x), y = 0$ in $\mathbb{R}^3 = \{x, y, z\}$. We can rotate this graph around the z axis to obtain a surface of revolution. The surface is realized as the level set of the submersion $f(x, y, z) = z - F(\sqrt{x^2 + y^2})$. f satisfies the invariance $f(Aq) = f(q)$, where $A \in SO(2)$ and the action of $SO(2)$ on $\mathbb{R}^3$ is rotations about the vertical axis, as defined in Section 1.5. We suppose that a single particle of mass m moves on the surface under the influence of the gravitational potential $V = gz$, where g is a positive constant. We seek the relative equilibria, which are the positions and velocities so that the particle will travel at constant speed in circles parallel to the xy plane. See Figure 2.1.

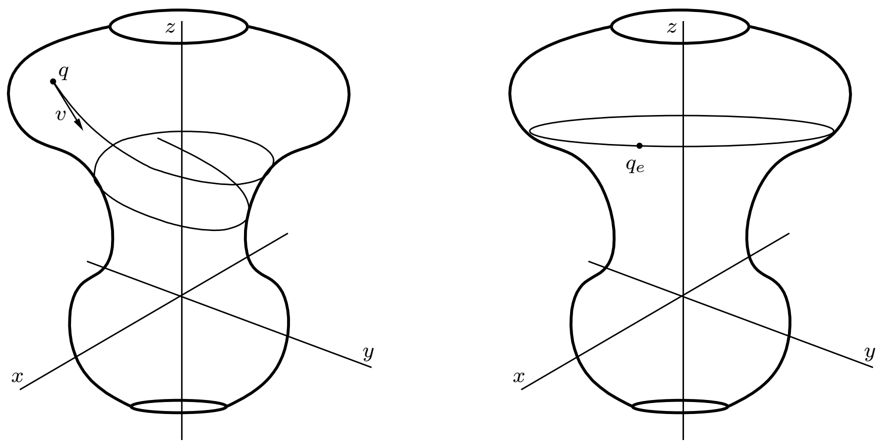

Figure 2.1: On the left, a particle with a generic initial position and velocity moves on a surface of revolution. On the right, the initial position and velocity are such that the particle moves in a circle in the horizontal plane with angular frequency ξ_e. The pair (q_e, ξ_e) is a relative equilibrium of this system.

The general elements are easily specified to the example. The configuration space q is $f^{-1}(0) \subseteq \mathbb{R}^3$. The kinetic energy metric in $\mathbb{R}^3$ is the usual metric scaled by the mass, so that

$$\langle v, w \rangle = mv \cdot w, \qquad M = m\mathbf{1}.$$

The Lagrangian is the function of $(\mathbb{R}^3)^2 = \{q, v\}$ defined by

$$L = \frac{m}{2}|v|^2 - V(q).$$

The group G is $SO(2)$ and the action on Q has already been defined. Since the infinitesimal generator is $I_\xi(q) = \xi\mathbf{k}^\wedge q = \xi\mathbf{k} \times q$ (equation (1.8)), the amended potential is easily computed to be

$$V_\xi = mgz - \frac{m}{2}\xi^2(x^2 + y^2).$$

From equation (2.17), the momentum is

$$J_\xi(q, v) = v^t M(\xi\mathbf{k} \times q) = \xi(Mv) \cdot (\mathbf{k} \times q) = \xi\mathbf{k} \cdot (q \times mv).$$

Thus, with the identification of $so(2)^*$ with $\mathbb{R}$, the momentum J takes values in $\mathbb{R}$

and is given by

$$J(q, v) = m\mathbf{k} \cdot q \times v$$

i.e. the physical angular momentum about the vertical axis.

Using (2.14), we are going to find the positions and the velocities of the particle that give horizontal, circular motion. We obtain

$$\nabla \left(\frac{m}{2} \xi^2 (x^2 + y^2) - mgz \right) = \lambda \nabla \left(z - F(\sqrt{x^2 + y^2}) \right).$$

We can rewrite this equation as the system

$$\begin{aligned}
m\xi^2 x &= -\lambda F'(\sqrt{x^2 + y^2}) \frac{x}{\sqrt{x^2 + y^2}} \\
m\xi^2 y &= -\lambda F'(\sqrt{x^2 + y^2}) \frac{y}{\sqrt{x^2 + y^2}} \\
mg &= -\lambda \\
F(\sqrt{x^2 + y^2}) &= z.
\end{aligned}$$

We may assume that $y = 0$ and $x > 0$. Therefore, three equations and three unknowns remain:

$$m\xi^2 x = -\lambda F'(x), \qquad mg = -\lambda, \qquad F(x) = z.$$

Thus $\lambda = mg$ and the first equation gives

$$\xi^2 x = gF'(x)$$

and given ξ we can find x from here. At the end, we find z from the third equation.

Immediately we can observe that there are no relative equilibria in the part of the domain of F for which F' is negative. This is physically natural since, for a circular motion about the vertical axis, the surface must maintain a constraint force that acts inwards towards that axis. At places where F' is negative this force acts downwards and cannot balance the motion against the downward influence of gravity.

Let the surface of revolution be the sphere of radius r. We are looking for the positions and velocities we have to give a particle so that it would travel in circles parallel to the rim. We could cast this as a special case of the above development, but instead we choose to use the equation

$$\nabla V_{\xi_e} = \lambda \nabla f,$$

where

$$V_{\xi_e}(x,y,z) = mgz - \frac{m}{2}\xi^2(x^2+y^2), \qquad f(x,y,z) = x^2+y^2+z^2-r^2.$$

We obtain

$$m\xi^2 x = -2x\lambda, \qquad m\xi^2 y = -2y\lambda, \qquad mg = 2z\lambda, \qquad x^2+y^2+z^2 = r^2.$$

We can assume that $y = 0$ and $x \geq 0$. This gives us the system

$$m\xi^2 x = -2x\lambda, \qquad mg = 2z\lambda, \qquad x^2+z^2 = r^2.$$

If $x = 0$, then $z = -r$, so the particle will stay at the lowest point on the sphere. Otherwise, $\lambda = -m\xi^2/2$. The particle can have arbitrary angular velocity ξ. We obtain

$$z = \frac{mg}{2\lambda} = -\frac{mg}{m\xi^2} = -\frac{g}{\xi^2}$$

and

$$x = \sqrt{r^2 - z^2} = \sqrt{r^2 - \frac{g^2}{\xi^4}}.$$

Thus, there is a one parameter family of relative equilibria, parameterized by the variable ξ.

Chapter 3

Single Rigid Body

3.1 $SO(3)$

$SO(3)$ is the set of 3×3 matrices A such that $\det A > 0$ and $A^t A = \mathbf{1}$. This is a Lie group in the sense of this thesis: the map $h(A) = A^t A$ defined on $\det A > 0$ is a submersion by an analogous argument to the one given for the Lie group $SO(2)$ in Section 1.5.

As in Section 1.5, the derivative of h at $\mathbf{1}$ is $\dot{A} \mapsto \dot{A} + \dot{A}^t$, and so the Lie algebra $so(3)$ of $SO(3)$ is the 3 dimensional vector space of 3×3 antisymmetric matrices. We identify $so(3)$ with $\mathbb{R}^3$ by the linear isomorphism $a \mapsto a^\wedge$ defined by

$$a^\wedge = \begin{bmatrix} 0 & -a^3 & a^2 \\ a^3 & 0 & -a^1 \\ -a^2 & a^1 & 0 \end{bmatrix}.$$

$a^\wedge x = a \times x$ for all $x \in \mathbb{R}^3$ and if $A \in SO(3)$ then $A(a \times x) = Aa \times Ax$ so

$$(Aa)^\wedge x = (Aa) \times x = A(a \times A^t x) = Aa^\wedge A^t x.$$

Since x was arbitrary, we have $Aa^\wedge A^{-1} = (Aa)^\wedge$, which means that the adjoint action under the identification of $so(3)$ with $\mathbb{R}^3$ is ordinary matrix multiplication i.e. $\mathrm{Ad}_A \xi = A\xi$. We will identify the dual $so(3)^*$ with $\mathbb{R}^3$ using the ordinary dot product, so that $\mu \in \mathbb{R}^3 \equiv so(3)^*$ acts on $\xi \in \mathbb{R}^3 \equiv so(3)$ by $\langle \mu, \xi \rangle = \mu \cdot \xi$. With this identification

$$\langle \operatorname{CoAd}_A \mu, \xi \rangle = \langle \mu, \operatorname{Ad}_{A^{-1}} \xi \rangle = \mu \cdot (A^t \xi) = A\mu \cdot \xi$$

and since ξ is arbitrary we have $\operatorname{CoAd}_A \mu = A\mu$ i.e. the coadjoint action is also ordinary matrix multiplication.

Since

$$\mathbf{k}^\wedge = \begin{bmatrix} \mathbf{J} & 0 \\ 0 & 0 \end{bmatrix}$$

where $\mathbf{J}$ is the matrix (1.7), and since we have, in view of (1.6),

$$\exp(\mathbf{k}\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

so $\exp(\mathbf{k}\theta)$ is the rotation matrix about $\mathbf{k}$ through θ radians, using the right hand rule. If $\xi \in so(3)$ then there is a rotation matrix $A \in SO(3)$ such that $|\xi|A\mathbf{k} = \xi$. But then

$$\exp(\xi t) = \exp(A\mathbf{k}|\xi|t) = \exp(A\mathbf{k}^\wedge A^t|\xi|t) = A\exp(\mathbf{k}|\xi|t)A^t,$$

so, with respect to the basis $A\mathbf{i}, A\mathbf{j}, A\mathbf{k}$, $\exp(\xi t)$ is the rotation about the axis $A\mathbf{k}$ through $|\xi|$ radians, using the right hand rule. The axis $A\mathbf{k}$ is the one along ξ. Consequently, $\exp\xi$ is the rotation about the axis along ξ through $|\xi|$, radians, using the right hand rule.

3.2 Lagrangian

We consider a rigid body specified by a mass density $\rho(x)$ at rest in an inertial reference frame with coordinates

$$x = \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}.$$

We assume that the center of mass of the body is at $x = 0$. A general configuration of the body corresponds to a rotation matrix $A \in SO(3)$: in the configuration A the point x in the reference frame is located at Ax. We assume that no external forces act on the body, so for this system, there is no potential energy and the Lagrangian is the system's kinetic energy.

If $A(t) \in SO(3)$ is a time dependent configuration, and $\dot{A}(t)$ denotes the derivative of $A(t)$ with respect to t, then the kinetic energy, and so the Lagrangian, is

$$\begin{aligned} L &= \frac{1}{2} \int \rho(x) |\dot{A}x|^2 \, dV \\ &= \frac{1}{2} \int \rho(x) \operatorname{trace}\big((\dot{A}x)(\dot{A}x)^t\big) \, dV \\ &= \frac{1}{2} \operatorname{trace}\left(\dot{A} \left[\int x x^t \rho(x) \, dV \right] \dot{A}^t \right). \end{aligned}$$

The *coefficient of inertia* of the body is defined as

$$\mathbb{I} = \int \rho(x) x x^t \, dV$$

and with this definition, the Lagrangian is

$$L(A, \dot{A}) = \frac{1}{2} \operatorname{trace}(\dot{A} \mathbb{I} \dot{A}^t).$$

This is of the form kinetic minus potential, with the kinetic energy given by the metric

$$\langle \dot{A}_1, \dot{A}_2 \rangle = \operatorname{trace}(\dot{A}_1 \mathbb{I} \dot{A}_2) \tag{3.1}$$

and zero potential energy

Up to this point, $SO(3)$ plays the role of the configuration space of the system rather than as a symmetry group. Now, we will see that, in this system, it also plays the role of the symmetry group. The *left action* of $SO(3)$ on itself is defined by $(B, A) \mapsto BA$. The metric (3.1) is invariant, since

$$\begin{aligned} \langle B\dot{A}_1, B\dot{A}_2 \rangle &= \frac{1}{2} \operatorname{trace}\big((B\dot{A}_1) I (B\dot{A}_2)^t\big) \\ &= \frac{1}{2} \operatorname{trace}\big(B(\dot{A}_1 I \dot{A}_2{}^t) B^t\big) \\ &= \frac{1}{2} \operatorname{trace}\big(B^t B(\dot{A}_1 I \dot{A}_2{}^t)\big) \\ &= \langle \dot{A}_1, \dot{A}_2 \rangle. \end{aligned}$$

Thus, the single rigid body is a Lagrangian system with symmetry as in Section 2.2. The configuration space is $SO(3) = \{A\}$ and symmetry group is $SO(3) = \{B\}$.

Pairs $(A, \dot{A}) \in TSO(3)$ are those for which the 3×3 matrix $A^t\dot{A}$ is antisymmetric. It is common and convenient to work with pairs (A, Ω) rather that $(A, \dot{A})$, where $\Omega \in \mathbb{R}^3$ is defined by

$$\Omega^\wedge = A^t\dot{A}.$$

Ω, called the *body angular velocity* has the physical interpretation as the angular velocity of the body written instantaneously in the basis $A\mathbf{i}, A\mathbf{j}, A\mathbf{k}$. To see this, consider a time dependent state of configurations $A(t)$. In those configurations, the point of the body with coordinates x moves along the curve $A(t)x$ and has instantaneous velocity at time t of

$$A'(t)x = A(t)\big(A(t)^tA'(t)\big)x = A(t)(\Omega \times x) = A(t)\Omega \times A(t)x, \tag{3.2}$$

where $\Omega^\wedge = A(t)^tA'(t)$. The point of the body which in the reference configuration was at x is at $\tilde{x} = A(t)x$ at time t. Since equation (3.2) gives its velocity as $A(t)\Omega \times \tilde{x}$, we read that it has angular velocity $A(t)\Omega$. Writing this in the basis $A\mathbf{i}, A\mathbf{j}, A\mathbf{k}$ gives the vector $A(t)^tA(t)\Omega = \Omega$, as required.

The kinetic energy metric, when it is evaluated on elements $\dot{A}_1, \dot{A}_2$ such that $(A, \dot{A}_1), (A, \dot{A}_2) \in TSO(3)$, may be written in terms of Ω_1, Ω_2, where $\Omega_1^\wedge = A^t\dot{A}_1$ and $\Omega_2^\wedge = A^t\dot{A}_2$, as follows:

$$\begin{aligned}
\langle \dot{A}_1, \dot{A}_2 \rangle &= \operatorname{trace}(\dot{A}_1\mathbb{I}\dot{A}_2{}^t) \\
&= \operatorname{trace}(A^t\dot{A}_1\mathbb{I}\dot{A}_2{}^tA) \\
&= \operatorname{trace}\big(\Omega_1{}^\wedge\mathbb{I}(\Omega_2{}^\wedge)^t\big) \\
&= \operatorname{trace}\big(-\Omega_2{}^\wedge\Omega_1{}^\wedge)\mathbb{I}\big).
\end{aligned}$$

Note that $\Omega_2{}^\wedge\Omega_1{}^\wedge = \Omega_1\Omega_2{}^t - \Omega_1 \cdot \Omega_2\mathbf{1}$, so

$$\begin{aligned}
\operatorname{trace}\big(-\Omega_2{}^\wedge\Omega_1{}^\wedge\mathbb{I}) &= \operatorname{trace}\big((\Omega_1 \cdot \Omega_2)\mathbb{I} - \Omega_1\Omega_2{}^t\mathbb{I}\big) \\
&= \Omega_1{}^t\left(\operatorname{trace}(\mathbb{I})\mathbf{1}\right)\Omega_2 - \Omega_1{}^t\mathbb{I}\Omega_2 \\
&= \Omega_1{}^t\big(\operatorname{trace}(\mathbb{I})\mathbf{1} - \mathbb{I}\big)\Omega_2 \\
&= \Omega_1{}^t\mathbb{J}\Omega_2,
\end{aligned}$$

where $\mathbb{J}$ is the *moment of inertia matrix*, defined by

$$\mathbb{J} = \mathrm{trace}(\mathbb{I})\mathbf{1} - \mathbb{I} = \begin{bmatrix} I^{22}+I^{33} & -I^{12} & -I^{13} \\ -I^{21} & I^{11}+I^{33} & -I^{23} \\ -I^{31} & -I^{32} & I^{11}+I^{22} \end{bmatrix}.$$

Thus,

$$\langle \dot{A}_1, \dot{A}_2 \rangle = {\Omega_1}^t \mathbb{J} \Omega_2,$$

whenever $(A, \dot{A}_1), (A, \dot{A}_2) \in TSO(3)$ as long as $\Omega_1^\wedge = A^t \dot{A}_1, \Omega_2^\wedge = A^t \dot{A}_2$, and

$$L(A, \Omega) = \frac{1}{2} \Omega^t \mathbb{J} \Omega.$$

The infinitesimal generator of the left action of $SO(3)$ on itself is

$$I_\xi A = \left.\frac{d}{dt}\right|_{t=0} \exp(\xi t) A = \xi^\wedge A.$$

If $\xi \in so(3)$ then the momentum J_ξ is, from equation (2.17),

$$J_\xi(A, \dot{A}) = \langle \dot{A}, \xi^\wedge A \rangle.$$

Since $A^t \dot{A} = \Omega^\wedge$ and $A^t \xi^\wedge A = (A^t \xi)^\wedge$, the momentum in terms of the matrix $\mathbb{J}$ is

$$J_\xi(A, \Omega) = \Omega^t \mathbb{J}(A^t \xi) = \xi \cdot A \mathbb{J} \Omega.$$

With the identification of $so(3)^*$ with $\mathbb{R}^3$, this is

$$J = A \mathbb{J} \Omega.$$

3.3 Relative Equilibria

Given any $A \in SO(3)$ we can find a B such that $BA = \mathbf{1}$, namely $B = A^{-1}$. So every element of $Q = SO(3)$ is in the group orbit of $\mathbf{1}$. Thus we can assume $A = \mathbf{1}$ when finding the relative equilibria.

The amended potential for this system is

$$V_\xi(A) = -\frac{1}{2}(A^t \xi)^t \mathbb{J}(A^t \xi),$$

where $\xi \in so(3)$. Using equation (2.13), $(\mathbf{1}, \xi)$ is a relative equilibrium if and only if, for all $w \in \mathbb{R}^3$,

$$0 = -\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \frac{1}{2}\big((\mathbf{1}+\epsilon w^\wedge)^t \xi\big)^t \mathbb{J}((\mathbf{1}+\epsilon w^\wedge)^t \xi)$$

$$
\begin{aligned}
&= -\xi^t \mathbb{J}\big(w^\wedge \xi\big) \\
&= -(\mathbb{J}\xi) \cdot (w \times \xi) \\
&= w \cdot \big((\mathbb{J}\xi) \times \xi\big).
\end{aligned}
$$

This is equivalent to $\mathbb{J}\xi \times \xi = 0$ i.e. ξ and $\mathbb{J}\xi$ are parallel. Thus, $(\mathbf{1}, \xi)$ is a relative equilibrium if and only if $\xi = 0$ or ξ is an eigenvector of $\mathbb{J}$, and there are three cases:

1. $\mathbb{J}$ has three distinct eigenvalues. Then a choice of three linearly independent eigenvectors a_1, a_2, a_3 gives the three one parameter families of relative equilibria $(\mathbf{1}, \xi)$ where $\xi = ta_i$, $t \neq 0$. These correspond to steady rotations of the body about the three perpendicular axes defined by the eigenvectors of $\mathbb{J}$. These axes are called the principle axes of inertia. There is also the solution $\xi = 0$ corresponding to the situation where the body is still.

2. $\mathbb{J}$ has two distinct eigenvalues. Then there is one eigenspace of dimension 1, say spanned by the eigenvector a_1, and a perpendicular eigenspace of dimension 2, say spanned by a_2, a_3. There is a one parameter class of relative equilibria $\xi = ta_1$, $t \neq 0$ and a two parameter class of relative equilibria $\xi = t_1 a_1 + t_2 a_2$, t_1, t_2 not both zero.

   This case is obtained from a rigid body with a mass distribution ρ which is symmetric about some axis. For example, if ρ is symmetric about the $\mathbf{k}$ axis then $\mathbb{I}$ and $\mathbb{J}$ are both diagonal with $I^{11} = I^{22}$ and $J^{11} = J^{22}$. In this case there is a one parameter class of relative equilibria corresponding to rotations about the symmetry axis and a two parameter class of relative equilibria corresponding to rotations about any axis perpendicular to the symmetry axis.

3. $\mathbb{J}$ has all eigenvalues equal. Any $\xi \in \mathbb{R}^3$ is a solution.

   This case is obtained from a rigid body with a spherically symmetric mass distribution, where $\mathbb{I}$ and $\mathbb{J}$ are both constant multiples of the identity.

Generically $\mathbb{J}$ has three distinct eigenvalues, corresponding to case 1. We further discuss systems with axially symmetric mass densities in the next section.

3.4 Axially symmetric mass distributions

Suppose that the mass density ρ in the reference frame has an axial symmetry about the $\mathbf{k}$ axis. Then $\mathbb{I}$ is diagonal and $I^{11} = I^{22}$. This implies $\mathbb{J}$ is also diagonal with $J^{11} = J^{22}$.

A circular mass distribution gives rise to a Lagrangian system with a larger symmetry group, in comparison with a generic mass distribution. Indeed, we can consider the action of $\tilde{B} \in SO(2)$ on $A \in SO(3)$ by $A \mapsto A\tilde{B}_3{}^t$ where

$$\tilde{B}_3 = \begin{bmatrix} \tilde{B} & 0 \\ 0 & 1 \end{bmatrix}.$$

This is called the *action of* $SO(2)$ *on* $SO(3)$ *by right multiplication by the inverse.*

As is easily verified, $\mathbb{I}$ and $\mathbb{J}$ commute with the matrices of the form B_3, as long as $I^{11} = I^{22}$ (or equivalently $J^{11} = J^{22}$):

$$\begin{bmatrix} I^{11}\mathbf{1} & 0 \\ 0 & I^{33} \end{bmatrix} \begin{bmatrix} \tilde{B} & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \tilde{B} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I^{11}\mathbf{1} & 0 \\ 0 & I^{33} \end{bmatrix}.$$

It follows that the metric is invariant under this new $SO(2)$ action since

$$\begin{aligned} \langle \dot{A}_1\tilde{B}_3, \dot{A}_2\tilde{B}_3 \rangle &= \operatorname{trace}\big((\dot{A}_1\tilde{B}_3)\mathbb{I}(\dot{A}_2\tilde{B}_3)^t\big) \\ &= \operatorname{trace}\big(\dot{A}_1\tilde{B}_3\mathbb{I}\tilde{B}_3{}^t\dot{A}_2{}^t\big) \\ &= \operatorname{trace}\big(\dot{A}_1\mathbb{I}\tilde{B}_3\tilde{B}_3{}^t\dot{A}_2{}^t\big) \\ &= \operatorname{trace}\big(\dot{A}_1\mathbb{I}\dot{A}_2{}^t\big) = \langle \dot{A}_1, \dot{A}_2 \rangle. \end{aligned}$$

We see if the mass density has a circular symmetry then the symmetry group of the Lagrangian is $SO(3) \times SO(2)$ where the first factor acts on the configuration space $SO(3)$ on the left while the second factor acts by right multiplication by the inverse.

$SO(3) \times SO(2) \subset \mathrm{GL}(3) \times \mathrm{GL}(2)$ is a Lie group in the sense of this thesis by using the product of the submersions which define the groups $SO(3)$ and $SO(2)$. The Lie algebra of $SO(3) \times SO(2)$ is thus the direct product $so(3) \times so(2)$ of the 3×3 antisymmetric matrices and the 2×2 antisymmetric matrices, and thus may be identified with $\mathbb{R}^3 \times \mathbb{R} = \big\{(\xi, \sigma)\big\}$ by

$$(\xi, \sigma) \mapsto (\xi^{\wedge}, \sigma\mathbf{J}).$$

The dual of $so(3) \times so(2)$ is $\mathbb{R}^3 \times \mathbb{R} = \big\{(\mu, \tau)\big\}$ where (μ, τ) acts on $(\xi\sigma)$ by

$$\langle(\mu, \tau), (\xi\sigma)\rangle = \mu \cdot \xi + \tau\sigma.$$

The exponential map is $\exp(\xi, \sigma) = \big(\exp(\xi^\wedge), \exp(\sigma\mathbf{J})\big)$.

By definition of the action, $\exp(\xi, \sigma)A = \exp(\xi)A\exp(-\sigma\mathbf{k}^\wedge)$, so the infinitesimal generator of the action of $SO(3) \times SO(2)$ on $SO(3)$ is

$$I_{(\xi,\sigma)}A = \frac{d}{dt}\bigg|_{t=0} \exp(\xi^\wedge t)A\exp(-\mathbf{k}^\wedge\sigma) = \xi^\wedge A - \sigma A\mathbf{k}^\wedge$$

and the momentum is

$$J_{(\xi,\sigma)}(A, \dot{A}) = \langle\dot{A}, A\xi^\wedge A - \sigma A\mathbf{k}\rangle.$$

Since $A^t\dot{A} = \Omega^\wedge$ and $A^t(\xi^\wedge A - A\sigma\mathbf{k}^\wedge) = (A^t\xi - \sigma\mathbf{k})^\wedge$, the momentum in terms of $\mathbb{J}$ is

$$J_{(\xi,\sigma)}(A, \Omega) = \Omega^t\mathbb{J}(A^t\xi - \sigma\mathbf{k}) = \xi \cdot (A\mathbb{J}\Omega) - \sigma J^{33}(\mathbf{k} \cdot \Omega).$$

Thus, the momentum is

$$J = (A\mathbb{J}\Omega, -J^{33}\mathbf{k} \cdot \Omega).$$

The amended potential is

$$V_{(\xi,\sigma)}(A) = -\frac{1}{2}(A^t\xi^\wedge - \sigma\mathbf{k})^t\mathbb{J}(A^t\xi^\wedge - \sigma\mathbf{k})$$

so $(\mathbf{1}, \xi, \sigma)$ is a relative equilibrium if and only if, for all $w \in \mathbb{R}^3$,

$$\begin{aligned} 0 &= -\frac{d}{d\epsilon}\bigg|_{\epsilon=0} \frac{1}{2}\big((\mathbf{1} + \epsilon w)^\wedge\xi - \sigma\mathbf{k}\big)^t\mathbb{J}\big((\mathbf{1} + \epsilon w)^\wedge\xi - \sigma\mathbf{k}\big) \\ &= -(\xi - \sigma\mathbf{k})^t\mathbb{J}(w \times \xi) \\ &= -\big(\mathbb{J}(\xi - \sigma\mathbf{k})\big) \cdot (w \times \xi) \\ &= w \cdot \big(\big(\mathbb{J}(\xi - \sigma\mathbf{k})\big) \times \xi\big). \end{aligned}$$

This is equivalent to

$$\mathbb{J}(\xi - \sigma\mathbf{k})\big) \times \xi = \begin{bmatrix} (J^{11} - J^{33})\xi^2\xi^3 + J^{33}\xi^2\sigma \\ -(J^{11} - J^{33})\xi^1\xi^3 - J^{33}\xi^1\sigma \\ 0 \end{bmatrix} = 0$$

and so we seek Ω, ξ, σ such that

$$\begin{aligned}
&\xi^2\big((J^{11} - J^{33})\xi^3 + J^{33}\sigma\big) = 0,\\
&\xi^1\big((J^{11} - J^{33})\xi^3 + J^{33}\sigma\big) = 0,\\
&\Omega = \xi - \sigma\mathbf{k}.
\end{aligned}$$

From the third equation, $\Omega^1 = \xi^1$ and $\Omega^2 = \xi^2$. There are two cases:

1. The case where one of Ω^1 or Ω^2 is not zero. Then the equations are equivalent to

$$\xi^1 = \Omega^1, \quad \xi^2 = \Omega^2, \quad (J^{11} - J^{33})\xi^3 + J^{33}\sigma = 0, \quad \xi^3 - \sigma = \Omega^3.$$

   with solution

$$\xi^1 = \Omega^1, \quad \xi^2 = \Omega^2, \quad \xi^3 = \frac{J^{33}\Omega^3}{J^{11}}, \quad \sigma = \frac{(J^{33} - J^{11})\Omega^3}{J^{11}}.$$

2. The case where $\Omega^1 = \Omega^2 = 0$. Then the solution is

$$\xi^1 = 0, \quad \xi^2 = 0, \quad \xi^3 - \sigma = \Omega^3.$$

   The solution is not fully determined: one can take any ξ^3 and σ with difference equal to Ω^3. It does not matter which solution is chosen since the evolution is

$$\exp(\xi^3\mathbf{k})\exp(-\sigma\mathbf{k}) = \exp\big((\xi^3 - \sigma)\mathbf{k}\big),$$

   and this only depends on the difference $\xi^3 - \sigma$.

We see that this system is completely solved by its symmetry, in the sense that every $(\mathbf{1}, \Omega)$ is a relative equilibrium, and given any $(\mathbf{1}, \Omega)$ we can find the corresponding ξ and σ to obtain the evolution $A(t) = \exp(\xi^3\mathbf{k})\exp(-\sigma\mathbf{k})$. In these solutions, the rigid body rotates around the axis along ξ at angular velocity $|\xi|$ and spins on its own axis of symmetry at angular velocity σ.

Chapter 4

Coupled Underwater vehicles

4.1 $SE(3)$

A *Euclidean transformation* of $\mathbb{R}^3$ is a diffeomorphism of $\mathbb{R}^3$ of the form $x \mapsto Ax+a$ where $A \in SO(3)$ and $a \in \mathbb{R}^3$. If two Euclidean transformations are composed, so first $x \mapsto Bx + b$ and then $x \mapsto Ax + a$, the result is $x \mapsto A(Bx + b) + a = (AB)x + (a + Ab)$, which is another Euclidean transformation: rotation by AB and a translation by $a + Ab$. Thus the Euclidean transformations of $\mathbb{R}^3$ form a group.

Define the *Euclidean group* $SE(3)$ to be the set $SO(3) \times \mathbb{R}^3$, with group operations

$$1 = (\mathbf{1}, 0), \quad (A, a)(B, b) = (AB, a + Ab), \quad (A, a)^{-1} = (A^{-1}, -A^{-1}a).$$

When multiplication is so defined, the representation of $SE(3)$ on $\mathbb{R}^3$ by $(A, a)x = Ax+a$ satisfies $(A, a)\big((B, b)x\big) = \big((A, a)(B, b)\big)x$, and the group operation on $SE(3)$ mirrors the Euclidean transformations of $\mathbb{R}^3$. Indeed, the map which takes (A, a) to the Euclidean transformation $x \mapsto Ax + a$ is a group isomorphism from $SE(3)$ to the Euclidean transformations of $\mathbb{R}^3$.

$SE(3)$ may be viewed as a subgroup of the 4×4 matrices by identifying $(A, a) \in SE(3)$ with the matrix

$$\begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix}.$$

The matrix product of two such 4×4 matrices corresponds to the group product in $SE(3)$, as is easily verified. Define the map h on the 4×4 matrices with positive determinant (represented by 3×3, 3×1, 1×3 and 1×1 blocks) by

$$h \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = (A_{11}^t A_{11}, A_{21}, A_{22}).$$

Then $SE(3)$ is $h^{-1}(\mathbf{1}, 0, 1)$ and h is a submersion to the product of the the 3×3 symmetric matrices and $\mathbb{R}^4$, so $SE(3)$ is a Lie group in the sense of this thesis. The Lie algebra $se(3)$ of $SE(3)$ is the vector space of matrices

$$\begin{bmatrix} \dot{A}_{11} & \dot{A}_{12} \\ 0 & 0 \end{bmatrix}$$

where $\dot{A}_{11}$ is antisymmetric. This we identify with $\mathbb{R}^3 \times \mathbb{R}^3 = \{(\Omega, v)\}$ using

$$(\Omega, v) \mapsto \begin{bmatrix} \Omega^\wedge & v \\ 0 & 0 \end{bmatrix}.$$

As was done with $SO(3)$ in Section 3.2, tangent vectors of $SE(3)$ at (A, a) are identified with elements (A, a, Ω, v) using left translation, so that (A, a, Ω, v) means the tangent vector

$$\left.\frac{d}{dt}\right|_{t=0} (A, a)(\mathbf{1} + t\Omega^\wedge, tv) = (A, a, A\Omega^\wedge, Av).$$

With this understanding, the infinitesimal generator at (A, a) of $(\Omega, v) \in se(3)$ for the left action of $SE(3)$ on itself is

$$\left.\frac{d}{dt}\right|_{t=0} (A, a)^{-1}(\mathbf{1} + t\Omega^\wedge, tv)(A, a) = \big(A, a, A^{-1}\Omega, A^{-1}(\Omega \times a + v)\big), \tag{4.1}$$

while the infinitesimal generator $(\Omega, v) \in se(3)$ for the action of $SE(3)$ on itself by right multiplication by the inverse is

$$\left.\frac{d}{dt}\right|_{t=0} (A, a)^{-1}(A, a)(\mathbf{1} - t\Omega^\wedge, -tv) = (A, a, -\Omega, -v). \tag{4.2}$$

The exponential mapping of $SE(3)$ is best computed with the aid of the Euclidean transformations of $\mathbb{R}^3$. Suppose that $(\Omega, v) \in se(3)$ and we seek

$$\big(A(t), a(t)\big) = \exp(\Omega t, vt).$$

Note that

$$\frac{d}{dt}\begin{bmatrix} A(t) & a(t) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \Omega^\wedge & v \\ 0 & 0 \end{bmatrix}\begin{bmatrix} A(t) & a(t) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \Omega^\wedge A(t) & \Omega^\wedge a(t) + v \\ 0 & 0 \end{bmatrix},$$

so

$$\frac{dA}{dt} = \Omega^\wedge A, \qquad \frac{da}{dt} = \Omega^\wedge a + v.$$

Choose $x_0 \in \mathbb{R}^3$ and set $x(t) = A(t)x_0 + a(t)$. Then

$$\frac{dx}{dt} = \frac{dA}{dt}x_0 + \frac{da}{dt} = \Omega^\wedge Ax_0 + \Omega^\wedge a + v = \Omega^\wedge x + v = \Omega \times x + v.$$

so $x(t)$ satisfies the initial value problem

$$\frac{dx}{dt} = \Omega \times x + v, \qquad x(0) = x_0. \tag{4.3}$$

There are two cases. If $\Omega = 0$ then (4.3) has solution $x = vt + x_0$, which is the Euclidean transformation corresponding to the element $(0, vt) \in SE(3)$. If Ω is not zero, then find vectors $v^\parallel$ and $v^\perp$ parallel and perpendicular to Ω such that

$$v = v^\parallel - \Omega \times v^\perp.$$

This can be done explicitly using the identity

$$\Omega \times (\Omega \times v) = (\Omega \cdot v)\Omega - |\Omega|^2 v,$$

so

$$v = \frac{\Omega \cdot v}{|\Omega|^2}\Omega - \frac{1}{|\Omega|^2}\Omega \times (\Omega \times v)$$

i.e. one can set

$$v^\parallel = \frac{\Omega \cdot v}{|\Omega|^2}\Omega, \qquad v^\perp = \frac{1}{|\Omega|^2}\Omega \times v.$$

The initial value problem (4.3) becomes

$$\frac{dx}{dt} = \Omega \times (x - v^\perp) + v^\parallel, \qquad x(0) = x_0,$$

and setting $u = x - v^\perp$ gives

$$\frac{du}{dt} = \Omega \times u + v^\parallel.$$

A particular solution to this equation is $u = tv^\parallel$ and the general solution to the homogeneous equation is $u = \exp(\Omega^\wedge t)c$, where c is a constant. This gives

$$x = \exp(\Omega^\wedge t)(x_0 - v^\perp) + v^\perp + tv^\parallel$$

which is a Euclidean transformation corresponding to the element

$$\big(\exp(\Omega^\wedge t), \big(\exp(\mathbf{1} - \Omega^\wedge t)\big)v^\perp + tv^\parallel\big) \in SE(3).$$

To summarize, if $\Omega = 0$ then $\exp(\Omega t, vt)$ corresponds to the Euclidean transformation which translates points as though they all have velocity v. If $\Omega \neq 0$ then $\exp(\Omega t, vt)$ moves points of $\mathbb{R}^3$ in spirals around the axis $v^\perp + t\mathbb{R}\Omega$. This spiral is comprised of 1) a right hand rotation about Ω with angular velocity Ω, and 2) a translational velocity of $v^\parallel$, which is parallel to Ω.

The adjoint action of $SE(3)$ may be directly computed:

$$\begin{aligned}
\mathrm{Ad}_{(A,a)}(\Omega, v) &= \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Omega^\wedge & v \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A^{-1} & -A^{-1}a \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Omega^\wedge A^{-1} & -\Omega^\wedge A^{-1}a + v \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} A\Omega^\wedge A^{-1} & -A\Omega^\wedge A^{-1}a + Av \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} (A\Omega)^\wedge & -(A\Omega)^\wedge a + Av \\ 0 & 0 \end{bmatrix}
\end{aligned}$$

i.e.

$$\mathrm{Ad}_{(A,a)}(\Omega, v) = \big(A\Omega, -(A\Omega) \times a + Av\big).$$

We identify the dual $se(3)^*$ of $se(3)$ with $\big\{(\mu, p)\big\} = \mathbb{R}^3 \times \mathbb{R}^3$ using the rule

$$\langle(\mu, p), (\Omega, v)\rangle = \mu \cdot \Omega + p \cdot v.$$

For the coadjoint action, we compute

$$\begin{aligned}
\langle \mathrm{CoAd}_{A,a}(\mu, p), (\Omega, v)\rangle &= \langle(\mu, p), \mathrm{Ad}_{(A^{-1}, -A^{-1}a)}(\Omega, v)\rangle \\
&= \langle(\mu, p), (A^{-1}\Omega, (-(A^{-1}\Omega) \times (-A^{-1}a) + A^{-1}v)\rangle \\
&= \mu \cdot A^{-1}\Omega + p \cdot (A^{-1}\Omega \times A^{-1}a + A^{-1}v) \\
&= (A\mu) \cdot \Omega + (Ap) \cdot (\Omega \times a + v) \\
&= (A\mu + a \times (Ap)) \cdot \Omega + (Ap) \cdot v \\
&= \langle(A\mu + a \times (Ap), Ap), (\Omega, v)\rangle
\end{aligned}$$

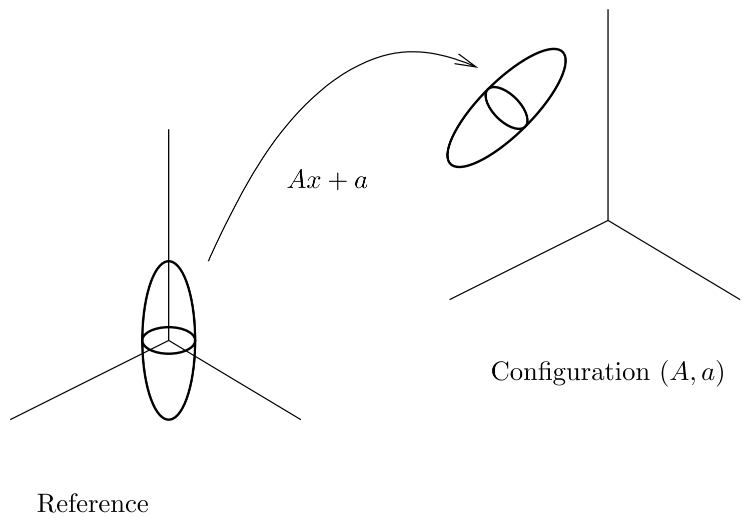

Figure 4.1: The configurations of a single underwater vehicle are parameterized by $(A, a) \in SE(3)$.

i.e.

$$\mathrm{CoAd}_{A,a}(\mu, p) = \big(A\mu + a \times (Ap), Ap\big).$$

4.2 Lagrangian formulation

The Lagrangian on $TSE(3)$ given by

$$L(R, a, \Omega, v) = \frac{1}{2}\Omega^t I \Omega + \Omega^t D v + \frac{1}{2} v^t M v + mgl(\mathbf{k} \cdot R\mathbf{k}),$$

approximates the motion of a neutrally buoyant vehicle submerged in an inviscid irrotational fluid [3][4][5]. In the model, the elements of $SE(3)$ parameterize the configurations of the vehicle by embedding a reference vehicle into the fluid. I, D, and M are constant 3×3 matrices that can be calculated from the shape and mass distribution of the vehicle. See Figure 4.1. The reference configuration is such that the center of buoyancy is at the origin and the center of mass is at distance l below the center of buoyancy. The vehicle is bottom heavy when $l > 0$ and top heavy when $l < 0$.

Two underwater vehicles may be modeled by using two copies of the system for

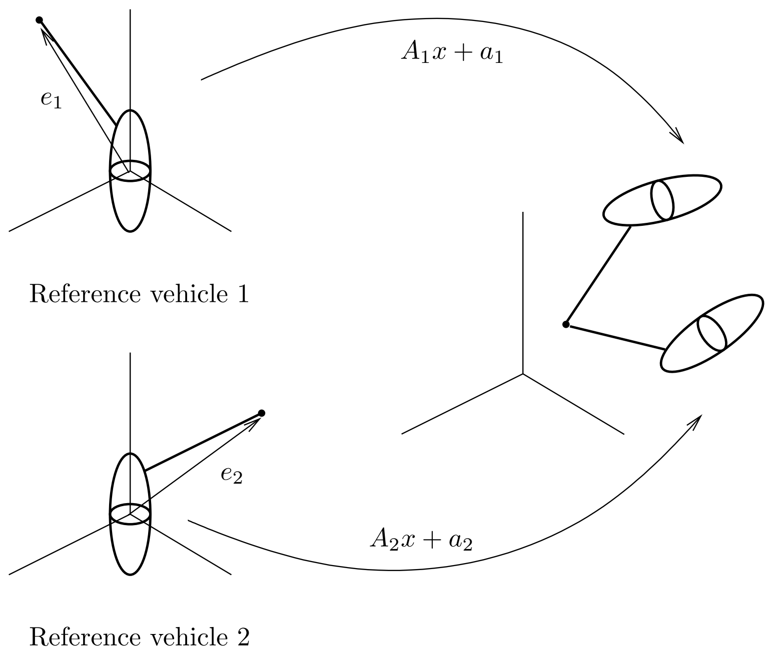

Figure 4.2: The configurations of two coupled underwater vehicles are parameterized by $(A_1, a_1, A_2, a_2) \in SE(3)^2$, subject to the coupling constraint $A_1e_1 + a_1 = A_2e_2 + a_2$.

a single underwater vehicle as above. See Figure 4.2. The simple direct product configuration space with Lagrangian L on $TSE(3)^2 = \{(A_i, a_i, \Omega_i, v_i)\}$ given by

$$L(A, a, \Omega, v) = \sum_{i=1}^{2} \frac{1}{2}\Omega_i^t I_i \Omega_i + \Omega_i^t D_i v_i + \frac{1}{2} v_i^t M_i v_i + mgl_i(\mathbf{k} \cdot A_i \mathbf{k})$$

models two noninteracting vehicles. To couple the vehicles, we assume that there is a spherically symmetric joint on vehicle i at location e_i in the reference configuration. The coupling is imposed by restricting to those configurations for which the joint on the first vehicle is located at the same place in space as the joint of the second vehicle. Thus the system has the constraint

$$A_1e_1 + a_1 = A_2e_2 + a_2.$$

We will assume $l = 0$, which implies $D = 0$, and that I and M are positive definite.

Were the coupling joint to distinguish the relative orientation of the vehicles then one would include a potential term to model this, but here we take the spherical coupling since it is the simplest possible. If a potential term of the form $A_1e_1 \cdot A_2e_2$ is included (i.e. subtract this from the Lagrangian) then there is a repulsive effect at the overlapped configuration and a potential well at the opposed configuration. One could also construct V corresponding to a potential well at some bent configuration. The coupling is purely mechanical and any hydrodynamic coupling is ignored.

4.3 Relative equilibria of coupled rigid bodies

Leaving the constraint explicit and using Lagrange multipliers leads to a different calculation than that used for finding the relative equilibria of coupled bodies in [7]. As well, the set-up of the coupled body system in [7] had the coupling joint in the reference system at the origin, whereas in the set-up above, it is not. So it is prudent to derive the equations for the relative equilibria in a set-up for the coupled body system which is analogous to the above, and make sure these equations correspond to those derived in [7].

The system of two coupled bodies moving freely in space has $M_1 = m_1\mathbf{1}$ and $M_2 = m_2\mathbf{1}$ (i.e. M_i is m_i times the 3×3 identity matrix). So

$$L = \frac{1}{2}\sum_{i=1}^{2}\left(\Omega_i^t I_i \Omega_i + \frac{m_i}{2}|v_i|^2\right)$$

with constraint $h = 0$, where

$$h = (A_2e_2 + a_2) - (A_1e_1 + a_1).$$

One has

$$V_{(\Omega,\sigma_1,\sigma_2)} = -L, \quad \text{where} \quad \Omega_i = A_i^t\Omega - \sigma_i\mathbf{k}, \quad v_i = A_i^t(\Omega \times a_i + v),$$

and to find the relative equilibria, one finds $(A_i, a_i) \in SE(3)$, $(\Omega, v) \in se(3)$, and $\sigma_1, \sigma_2 \in \mathbb{R}$, such that, for all $w_1, w_2 \in \mathbb{R}^3$

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0}\left(V_{(\Omega,\sigma_1,\sigma_2)} + \lambda \cdot h\right) = 0$$

after making the replacements

$$A_i \leftarrow A_i + \epsilon A_i w_i^{\wedge}, \quad a_i \leftarrow a_i + \epsilon z_i$$

into $V_{(\Omega,\sigma_1,\sigma_2)} + \lambda \cdot h$. Differentiating at $\epsilon = 0$:

$$\begin{aligned}
&\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} V_{(\Omega,\sigma_1,\sigma_2)} \\
&\quad = -\big(\Omega_1' \cdot I_1\Omega_1 + \frac{1}{2}m_1(v_1 \cdot v_1)' + (1 \leftrightarrow 2)\big) \\
&\quad = -\big(\big(-w_1^\wedge A_1^t\Omega \cdot I_1\Omega_1 + m_1(\Omega \times z_1) \cdot (\Omega \times a_1 + v)\big) + (1 \leftrightarrow 2)\big) \\
&\quad = w_1 \times (A_1^t\Omega) \cdot I_1\Omega_1 - m_1(\Omega \times z_1) \cdot (\Omega \times a_1 + v)\big) + (1 \leftrightarrow 2)\big) \\
&\quad = A_1 w_1 \cdot (\Omega \times A_1 I_1 \Omega_1) - m_1 z_1 \cdot \big((\Omega \times a_1 + v) \times \Omega\big) + (1 \leftrightarrow 2)
\end{aligned}$$

while

$$\begin{aligned}
&-\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \lambda \cdot h \\
&\quad = -\lambda \cdot (A_2 w_2^\wedge e_2 + z_2 - A_1 w_1^\wedge e_1 - z_1) \\
&\quad = -A_2 w_2 \cdot (A_2 e_2 \times \lambda) + A_1 w_1 \cdot (A_1 e_1 \times \lambda) - z_2 \cdot \lambda + z_1 \cdot \lambda.
\end{aligned}$$

Equating $dV_{(\Omega,\sigma_1,\sigma_2)}$ and $-d(\lambda \cdot h)$, gives the equations for the relative equilibria as

$$\Omega \times A_1 I_1 \Omega_1 = A_1 e_1 \times \lambda, \quad \Omega \times A_2 I_2 \Omega_2 = -A_2 e_2 \times \lambda \tag{4.4}$$

$$m_1 \Omega \times (\Omega \times a_1 + v) = \lambda, \quad m_2 \Omega \times (\Omega \times a_2 + v) = -\lambda \tag{4.5}$$

$$A_1 e_1 + a_1 = A_2 e_2 + a_2, \tag{4.6}$$

$$\Omega_1 = A_1^t \Omega - \sigma_1 \mathbf{k}, \quad \Omega_2 = A_2^t \Omega - \sigma_2 \mathbf{k}. \tag{4.7}$$

A solution to these equations gives a solution to the system where the bodies have initial locations (A_1, a_i), have spin σ_i on their axes of symmetry, translate with velocity v, and rotate with angular velocity Ω.

We want to compare these equations with the ones in [7]. We can assume that the center of mass is at the origin, so that $m_1 a_1 + m_2 a_2 = 0$, and so from (4.6),

$$m_1 a_1 + m_2 a_2 = 0, \quad a_1 - a_2 = A_2 e_2 - A_1 e_1.$$

Solving these for a_1 and a_2 gives

$$a_1 = \frac{m_2}{m_1 + m_2}(A_2 e_2 - A_1 e_1), \quad a_2 = \frac{m_1}{m_1 + m_2}(A_1 e_1 - A_2 e_2). \tag{4.8}$$

Adding the two equations (4.5) gives $(m_1 + m_2) v \times \Omega = 0$ (i.e. v is arbitrary if $\Omega = 0$ and a constant of Ω if $\Omega \neq 0$), while the first of (4.8) into the first of (4.5) gives

$$\lambda = \epsilon\Omega \times \big(\Omega \times (A_2e_2 - A_1e_1)\big), \quad \epsilon = m_1m_2/(m_1+m_2).$$

This equation, $v \times \Omega = 0$, and equations (4.8) solve (4.5) and (4.6) in the case $m_1a_1 + m_2a_2 = 0$, and so only (4.4) and (4.7) remain. The right side of the first equation of (4.4) is

$$\begin{aligned} A_1e_1 \times \lambda &= \epsilon A_1e_1 \times \big(\Omega \times \big(\Omega \times (A_2e_2 - A_1e_1)\big)\big) \\ &= \epsilon\big(A_1e_1 \cdot \big(\Omega \times (A_2e_2 - A_1e_1)\big)\big)\Omega - (\epsilon A_1e_1 \cdot \Omega)\big(\Omega \times (A_2e_2 - A_1e_1)\big) \end{aligned}$$

and so equation (4.4) becomes

$$\Omega \times \big(A_1I_1\Omega_1 + \epsilon(A_1e_1 \cdot \Omega)(A_2e_2 - A_1e_1)\big) = \epsilon\big(A_1e_1 \cdot \big(\Omega \times (A_2e_2 - A_1e_1)\big)\big)\Omega. \quad (4.9)$$

Because the cross product is orthogonal to both of its arguments, this implies both sides vanish separately. The right side of (4.9) gives

$$\Omega \cdot A_1e_1 \times A_2e_2 = -A_1e_1 \cdot \big(\Omega \times (A_2e_2 - A_1e_1)\big) = 0$$

which is a fact of the coupled body system that has been noted in [7]: at a relative equilibrium the angular velocity and the axes of the bodies are collinear. Setting

$$\tilde{I}_1 = I_1 - \epsilon(e_1^\wedge)^2 = I_1 - \epsilon e_1e_1^t + \epsilon|e_1|^2,$$

the left side of (4.9) gives

$$\begin{aligned} &\Omega \times \big(A_1I_1\Omega_1 + \epsilon(A_1e_1 \cdot \Omega)(A_2e_2 - A_1e_1)\big) \\ &\quad = \Omega \times \big(A_1I_1\Omega_1 - \epsilon(A_1e_1 \cdot \Omega)A_1e_1 + \epsilon(A_1e_1 \cdot \Omega)A_2e_2\big) \\ &\quad = \Omega \times \Big(A_1I_1\Omega_1 - \epsilon\big(e_1 \cdot (\Omega_1 + \sigma_1\mathbf{k})\big)A_1e_1 + \epsilon(A_1e_1 \cdot \Omega)A_2e_2\Big) \\ &\quad = \Omega \times \Big(A_1(I_1\Omega_1 - \epsilon e_1e_1^t\Omega_1 + \epsilon|e_1|^2\Omega_1) \\ &\qquad - A_1\big(\epsilon|e_1|^2\Omega_1 + \epsilon\sigma_1(e_1 \cdot \mathbf{k})e_1\big) + \epsilon(A_1e_1 \cdot \Omega)A_2e_2\Big) \\ &\quad = \Omega \times \Big(A_1\tilde{I}_1\Omega_1 - \epsilon|e_1|^2A_1(\Omega_1 + \sigma_1\mathbf{k}) + \epsilon(A_1e_1 \cdot \Omega)A_2e_2\Big) \\ &\quad = \Omega \times \Big(A_1\tilde{I}_1\Omega_1 - \epsilon|e_1|^2\Omega + \epsilon(A_1e_1 \cdot \Omega)A_2e_2\Big) \\ &\quad = \Omega \times \Big(A_1\tilde{I}_1\Omega_1 + \epsilon(A_1e_1 \cdot \Omega)A_2e_2\Big) \\ &\quad = 0. \end{aligned}$$

In the same way one get this equation, but with the indices 1 and 2 interchanged, from the second of equations (4.4). These two equations replicate the equations of [7], so the verification is complete.

4.4 Relative equilibria equations of coupled underwater vehicles

The Lagrangian for the coupled coupled underwater vehicle system differs from the coupled body system considered in the previous section only through the replacement of the scalars m_i by the diagonal matrices M_i. The change in the Lagrangian, due to the presence of the fluid, is just the replacement of $\frac{1}{2}m_i|v_i|^2$ with $\frac{1}{2}v_i^t M_i v_i$, and so only the effect of these terms must be recomputed. Proceeding as before,

$$\begin{aligned}
&-d\left(\frac{1}{2}v_1 \cdot M_1 v_1\right)(A_1, a_1, A_2, a_2)(w_1, z_1) \\
&\quad= -\left(\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} A_1^t(\Omega \times a_1 + v)\right) \cdot M_1 v_1 \\
&\quad= -\big(-w_1 \times A_1^t(\Omega \times a_1 + v) + A_1^t(\Omega \times z_1)\big) \cdot M_1 v_1 \\
&\quad= A_1 w_1 \cdot A_1\big(v_1 \times M_1 v_1\big) + z_1 \cdot (\Omega \times A_1 M_1 v_1),
\end{aligned}$$

so that

$$\begin{aligned}
&dV_{(\Omega,\sigma_1,\sigma_2)}(A_1, a_1, A_2, a_2)(w_1, w_2, z_1, z_2) \\
&\quad= \Big(A_1 w_1 \cdot (\Omega \times A_1 I_1 \Omega_1) + A_1 w_1 \cdot A_1(v_1 \times A_1 M_1 v_1) \\
&\qquad + z_1 \cdot (\Omega \times A_1 M_1 v_1)\Big) + (1 \leftrightarrow 2),
\end{aligned}$$

and, as before, this is to be equated, for all w_1, w_2, z_1, z_2, with

$$\begin{aligned}
&-\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \lambda \cdot h \\
&\quad= -\lambda \cdot (A_2 w_2^{\wedge} e_2 + z_2 - A_1 w_1^{\wedge} e_1 - z_1) \\
&\quad= -A_2 w_2 \cdot (A_2 e_2 \times \lambda) + A_1 w_1 \cdot (A_1 e_1 \times \lambda) - z_2 \cdot \lambda + z_1 \cdot \lambda.
\end{aligned}$$

The resulting equations for the relative equilibria are

$$\Omega \times A_1 I_1 \Omega_1 + A_1(v_1 \times M_1 v_1) = A_1 e_1 \times \lambda \tag{4.10}$$

$$\Omega \times A_2 I_2 \Omega_2 + A_2(v_2 \times M_2 v_2) = -A_2 e_2 \times \lambda \tag{4.11}$$

$$\Omega \times (A_1 M_1 v_1) = \lambda, \quad \Omega \times (A_2 M_2 v_2) = -\lambda, \tag{4.12}$$

$$A_1 e_1 + a_1 = A_2 e_2 + a_2, \tag{4.13}$$

$$\Omega_1 = A_1^t \Omega - \sigma_1 \mathbf{k}, \quad \Omega_2 = A_2^t \Omega - \sigma_2 \mathbf{k}, \tag{4.14}$$

$$v_1 = A_1^t(\Omega \times a_1 + v), \quad v_2 = A_2^t(\Omega \times a_2 + v). \tag{4.15}$$

We seek all $A_1, a_1, A_2, a_1, \Omega, v, \sigma_1, \sigma_2$ that solve these equations.

To simplify, to make it possible for there to be multiple rotating parts, and to increase the symmetry of the system, we will assume

1. identical coupled vehicles: $I_1 = I_2 = I$, $M_1 = M_2 = M$;
2. each vehicle has coincident centers of mass and buoyancy: $l_1 = l_2 = 0$, $D_i = 0$;
3. each vehicle is an ellipsoid: I and M are diagonal, and have positive entries;
4. each vehicle has an axis of symmetry: $I^{11} = I^{22}$ and $M^{11} = M^{22}$;
5. the coupling joint is along the symmetry axis, and in the same place on each vehicle: $e_1 = e_2 = e\mathbf{k}$.

Under these assumptions the system is Lagrangian with

$$L(A, a, \Omega, v) = \sum_{i=1}^{2} \left(\frac{1}{2}\Omega_i^t I \Omega_i + \frac{1}{2} v_i^t M v_i \right)$$

and the constraint

$$eA_1\mathbf{k} + a_1 = eA_2\mathbf{k} + a_2. \tag{4.16}$$

The symmetry is the group $SE(3) \times SO(2)^2 = \left\{ \left(A, \exp(\theta_1 \mathbf{k}^\wedge), \exp(\theta_2 \mathbf{k}^\wedge)\right) \right\}$ which acts on configuration space $SE(3)^2$ by diagonal left multiplication by A and right componentwise multiplication by $\left(\exp(\theta_1 \mathbf{k}^\wedge)^{-1}, \exp(\theta_2 \mathbf{k}^\wedge)^{-1}\right)$. From (4.1) and (4.2), the infinitesimal generator of $(\Omega, v, \sigma_1, \sigma_2)$ at (A_1, a_1, A_2, a_2) is

$$\Omega_i = A_i^{-1}\Omega - \sigma_i \mathbf{k}, \quad v_i = A_i^{-1}(\Omega \times a_i + v),$$

so, using the notation, $J(A_i, a_i, \Omega_i, v_i) = (\mu, p, \mu_1, \mu_2)$,

$\mu \cdot \Omega + p \cdot v + \mu_1 \sigma_1 + \mu_2 \sigma_2$

$$= \Omega_1 \cdot I\big(A_1^{-1}\Omega - \sigma_1 \mathbf{k}\big) + v_1 \cdot M\big(A_1^{-1}(\Omega \times a_1 + v)\big) + (1 \leftrightarrow 2)$$
$$= \Omega \cdot \big(A_1 I \Omega_1 + a_1 \times (A_1 M v_1)\big) + v \cdot (A_1 M v_1) - \sigma_1 I^{33} \mathbf{k} \cdot \Omega_1 + (1 \leftrightarrow 2).$$

Thus the momentum is

$$\mu = \sum_{i=1}^{2} A_i I \Omega_i + a_i \times (A_i M v_i), \quad p = \sum_{i=1}^{2} A_i M v_i, \tag{4.17}$$

$$\mu_1 = -I^{33}\mathbf{k} \cdot \Omega_1, \quad \mu_2 = -I^{33}\mathbf{k} \cdot \Omega_2. \tag{4.18}$$

We specialize equations (4.10)–(4.15) to the above assumptions as follows. For any three vector x and any 3×3 matrix C of the form

$$C = \begin{bmatrix} C^{11} & 0 & 0 \\ 0 & C^{11} & 0 \\ 0 & 0 & C^{33} \end{bmatrix}$$

there is the computation

$$\begin{aligned} x \times Cx &= x \times C\big((x - (x \cdot \mathbf{k})\mathbf{k}) + (x \cdot \mathbf{k})\mathbf{k}\big) \\ &= x \times \big(C^{11}\big(x - (x \cdot \mathbf{k})\mathbf{k}\big) + C^{33}(x \cdot \mathbf{k})\mathbf{k}\big) \\ &= \Delta C (x \cdot \mathbf{k}) x \times \mathbf{k}, \end{aligned}$$

where $\Delta C = C^{33} - C^{11}$. Also,

$$\mathbf{k} \times (Cx) = \mathbf{k} \times \big(C^{11}(x - (x \cdot \mathbf{k})\mathbf{k}) + C^{33}(x \cdot \mathbf{k})\mathbf{k}\big) = C^{11}\mathbf{k} \times x.$$

Multiplying (4.10) by A_1^t and substituting Ω from (4.14) gives

$$\begin{aligned} &(A_1^t \Omega) \times I_1 \Omega_1 + v_1 \times M_1 v_1 - e\mathbf{k} \times (A_1^t \lambda) \\ &= (\Omega_1 + \sigma_1 \mathbf{k}) \times I_1 \Omega_1 + v_1 \times M_1 v_1 - e\mathbf{k} \times (A_1^t \lambda) \\ &= \Omega_1 \times I_1 \Omega_1 + \sigma_1 \mathbf{k} \times I_1 \Omega_1 + v_1 \times M_1 v_1 - e\mathbf{k} \times (A_1^t \lambda) \\ &= \Delta I (\mathbf{k} \cdot \Omega_1) \Omega_1 \times \mathbf{k} + I^{11} \sigma_1 \mathbf{k} \times \Omega_1 + \Delta M (\mathbf{k} \cdot v_1) v_1 \times \mathbf{k} - e\mathbf{k} \times (A_1^t \lambda) \\ &= -\mathbf{k} \times \big(\big(\Delta I (\mathbf{k} \cdot \Omega_1) - I^{11}\sigma_1\big)\Omega_1 + \Delta M (\mathbf{k} \cdot v_1) v_1 + e A_1^t \lambda\big) \\ &= 0, \end{aligned}$$

which is equivalent to

$$\big(\Delta I (\mathbf{k} \cdot \Omega_1) - I^{11}\sigma_1\big)\Omega_1 + \Delta M (\mathbf{k} \cdot v_1) v_1 + e A_1^t \lambda = t_1 \mathbf{k}$$

for some $t_1 \in \mathbb{R}$. Multiplying by A_1 and using (4.14) gives

$$\begin{aligned}
A_1\Big(& \big(\Delta I(\mathbf{k}\cdot\Omega_1) - I^{11}\sigma_1\big)\Omega_1 + \Delta M(\mathbf{k}\cdot v_1)v_1 + eA_1^t\lambda - t_1\mathbf{k}\Big) \\
&= \big(\Delta I\,(\mathbf{k}\cdot A_1^t\Omega - \sigma_1) - I^{11}\sigma_1\big)(\Omega - \sigma_1 A_1\mathbf{k}) + \Delta M(\mathbf{k}\cdot v_1)A_1 v_1 + e\lambda - t_1 A_1\mathbf{k} \\
&= \big(\Delta I(A_1\mathbf{k})\cdot\Omega - I^{33}\sigma_1\big)(\Omega - \sigma_1 A_1\mathbf{k}) + \Delta M(\mathbf{k}\cdot v_1)A_1 v_1 + e\lambda - t_1 A_1\mathbf{k} \\
&= 0.
\end{aligned}$$

Also one obtains from (4.11) this same equation but with the subscripts 1 and 2 exchanged and the sign of λ changed. Thus equations (4.10) and (4.11) are reduced to

$$\begin{aligned}
&\big(\Delta I(A_1\mathbf{k})\cdot\Omega - I^{33}\sigma_1\big)(\Omega - \sigma_1 A_1\mathbf{k}) + \Delta M(\mathbf{k}\cdot v_1)A_1 v_1 - e\lambda - t_1 A_1\mathbf{k} = 0, \\
&\big(\Delta I(A_2\mathbf{k})\cdot\Omega - I^{33}\sigma_2\big)(\Omega - \sigma_2 A_2\mathbf{k}) + \Delta M(\mathbf{k}\cdot v_2)A_2 v_2 + e\lambda - t_2 A_2\mathbf{k} = 0.
\end{aligned} \tag{4.19}$$

Since t_1 and t_2 are arbitrary and they occur only in these two equations, we may absorb any multiple of $A_1\mathbf{k}$ into t_1 in the first equation and absorb any multiple of $A_2\mathbf{k}$ into t_2 in the second equation. So the equations (4.19) are equivalent to

$$\big(\Delta I(A_1\mathbf{k})\cdot\Omega - I^{33}\sigma_1\big)\Omega + \Delta M(\mathbf{k}\cdot v_1)A_1 v_1 - e\lambda - t_1 A_1\mathbf{k} = 0, \tag{4.20}$$

$$\big(\Delta I(A_2\mathbf{k})\cdot\Omega - I^{33}\sigma_2\big)\Omega + \Delta M(\mathbf{k}\cdot v_2)A_2 v_2 + e\lambda - t_2 A_2\mathbf{k} = 0. \tag{4.21}$$

Adding the two equations (4.12) gives

$$\Omega \times (A_1 M v_1 + A_2 M v_2) = 0$$

i.e. the possibility

$$A_1 M v_1 + A_2 M v_2 = t_3\Omega$$

for some $t_3 \in \mathbb{R}$, or the possibility that $\Omega = 0$. Equivalently, assuming $\Omega \neq 0$

$$\begin{aligned}
A_1 M v_1 + A_2 M v_2 &= A_1\big(M^{11}v_1 + \Delta M(\mathbf{k}\cdot v_1)\mathbf{k}\big) + A_2\big(M^{11}v_2 + \Delta M(\mathbf{k}\cdot v_2)\mathbf{k}\big) \\
&= M^{11}(A_1 v_1 + A_2 v_2) + \Delta M\big((\mathbf{k}\cdot v_1)A_1\mathbf{k} + (\mathbf{k}\cdot v_2)A_2\mathbf{k}\big) \\
&= t_3\Omega.
\end{aligned}$$

Without loss of generality, one may translate to the origin the joint connecting the bodies, so that

$$a_1 = -eA_1\mathbf{k}, \qquad a_2 = -eA_2\mathbf{k},$$

and from the first of equations (4.15),

$$\mathbf{k}\cdot v_1 = \mathbf{k}\cdot\big(-e(A_1^t\Omega)\times\mathbf{k} + A_1^t v\big) = (A_1\mathbf{k})\cdot v,$$

and similarly $\mathbf{k}\cdot v_2 = (A_2\mathbf{k})\cdot v$ from the second of (4.15), so

$$M^{11}(A_1v_1 + A_2v_2) + \Delta M\big((A_1\mathbf{k}\cdot v)A_1\mathbf{k} + (A_2\mathbf{k}\cdot v)A_2\mathbf{k}\big) = t_3\Omega.$$

Substituting equations (4.15) gives

$$\begin{aligned}
&M^{11}(A_1v_1 + A_2v_2) + \Delta M\big((A_1\mathbf{k}\cdot v)A_1\mathbf{k} + (A_2\mathbf{k}\cdot v)A_2\mathbf{k}\big)\\
&\quad = M^{11}(\Omega\times a_1 + v + \Omega\times a_2 + v) + \Delta M\big((A_1\mathbf{k}\cdot v)A_1\mathbf{k} + (A_2\mathbf{k}\cdot v)A_2\mathbf{k}\big)\\
&\quad = M^{11}\big(-e\Omega\times(A_1\mathbf{k}) + v - e\Omega\times(A_2\mathbf{k}) + v\big)\\
&\qquad\qquad + \Delta M\big((A_1\mathbf{k}\cdot v)A_1\mathbf{k} + (A_2\mathbf{k}\cdot v)A_2\mathbf{k}\big)\\
&\quad = \big(2M^{11}\mathbf{1} + \Delta M(A_1\mathbf{k}\mathbf{k}^tA_1^t + A_2\mathbf{k}\mathbf{k}^tA_2^t)\big)v - eM^{11}\Omega\times(A_1\mathbf{k} + A_2\mathbf{k})\\
&\quad = t_3\Omega,
\end{aligned}$$

which is equivalent (after rescaling t_3) to

$$\Big(\mathbf{1} + \frac{\Delta M}{2M^{11}}(A_1\mathbf{k}\mathbf{k}^tA_1^t + A_2\mathbf{k}\mathbf{k}^tA_2^t)\Big)v = t_3\Omega + \frac{e}{2}\Omega\times(A_1\mathbf{k} + A_2\mathbf{k}).$$

The fragment of (4.20) that depends on v_1 (substitute λ from Equation (4.12)) is

$$\begin{aligned}
&\Delta M(\mathbf{k}\cdot v_1)A_1v_1 + e\lambda\\
&\quad = \Delta M(\mathbf{k}\cdot v_1)\big(-e\Omega\times(A_1\mathbf{k}) + v\big) + e\Omega\times(A_1Mv_1)\\
&\quad = \Delta M(\mathbf{k}\cdot v_1)v + e\Omega\times A_1\big(-\Delta M(\mathbf{k}\cdot v_1)\mathbf{k} + Mv_1\big)\\
&\quad = \Delta M(\mathbf{k}\cdot v_1)v + eM^{11}\Omega\times A_1v_1\\
&\quad = \Delta M(\mathbf{k}\cdot v_1)v + eM^{11}\Omega\times\big(-e\Omega\times(A_1\mathbf{k}) + v\big)\\
&\quad = \Delta M(\mathbf{k}\cdot v_1)v - e^2M^{11}\big(\Omega\cdot(A_1\mathbf{k})\big)\Omega + eM^{11}\Omega\times v + e^2M^{11}|\Omega|^2A_1\mathbf{k}
\end{aligned}$$

so that after absorbing the multiple of $A_1\mathbf{k}$ into t_1, (4.20) becomes

$$\begin{aligned}
&\big(\Delta I(A_1\mathbf{k})\cdot\Omega - I^{33}\sigma_1\big)\Omega + \Delta M\big((A_1\mathbf{k})\cdot v\big)A_1v_1 + e\lambda\\
&\quad = \big(\Delta I(A_1\mathbf{k})\cdot\Omega - I^{33}\sigma_1\big)\Omega + \Delta M\big((A_1\mathbf{k})\cdot v)v\\
&\qquad\qquad - e^2M^{11}\big(\Omega\cdot(A_1\mathbf{k})\big)\Omega + eM^{11}\Omega\times v
\end{aligned}$$

$$= \big((\Delta I - e^2 M^{11})(A_1\mathbf{k})\cdot\Omega - I^{33}\sigma_1\big)\Omega + \Delta M(\mathbf{k}\cdot v_1)v + eM^{11}\Omega\times v$$
$$= t_1 A_1\mathbf{k}.$$

Hence equations (4.10)–(4.15) are reduced to

$$\begin{aligned}
&\big((\Delta I - e^2 M^{11})(A_1\mathbf{k})\cdot\Omega - I^{33}\sigma_1\big)\Omega \\
&\qquad + \Delta M\big((A_1\mathbf{k}\cdot v\big)v + eM^{11}\Omega\times v = t_1 A_1\mathbf{k}, && (4.22)\\
&\big((\Delta I - e^2 M^{11})(A_2\mathbf{k})\cdot\Omega - I^{33}\sigma_2\big)\Omega \\
&\qquad + \Delta M\big((A_2\mathbf{k})\cdot v\big)v + eM^{11}\Omega\times v = t_2 A_2\mathbf{k}, && (4.23)\\
&\Big(\mathbf{1} + \frac{\Delta M}{2M^{11}}(A_1\mathbf{k}\mathbf{k}^t A_1^t + A_2\mathbf{k}\mathbf{k}^t A_2^t)\Big)v = t_3\Omega + \frac{e}{2}\Omega\times(A_1\mathbf{k} + A_2\mathbf{k}). && (4.24)
\end{aligned}$$

Equation (4.24) has been derived under the assumption that $\Omega \neq 0$, and there is also the possibility that $\Omega = 0$ replaces equation (4.24). However,

$$\begin{aligned}
v^t\Big(\mathbf{1} + \frac{\Delta M}{2M^{11}}(A_1\mathbf{k}\mathbf{k}^t A_1^t + A_2\mathbf{k}\mathbf{k}^t A_2^t)\Big) &= |v|^2 + \frac{\Delta M}{2M^{11}}\big((\mathbf{k}\cdot A_1^t v)^2 + (\mathbf{k}\cdot A_1^t v)^2 \\
&\geq \left(1 + \frac{\Delta M}{M^{11}}\right)|v|^2 \\
&= \frac{M^{33}}{M^{11}}|v|^2,
\end{aligned}$$

so the matrix on the left in (4.24) is positive definite and hence nonsingular. Thus $\Omega = 0$ in (4.24) gives $v = 0$ i.e. the solution

$$\Omega = 0, \quad v = 0, \quad \text{all of } A_1, a_1, A_2, a_2, \sigma_1, \sigma_2 \text{ arbitrary.} \tag{4.25}$$

By substitution this is a solution to equations (4.22)–(4.24), so the possibility $\Omega = 0$ does not require separate consideration.

Chapter 5

Relative equilibria of coupled underwater vehicles

We will find a set of solutions of the equations

$$\big((\Delta I - e^2 M^{11})(A_1\mathbf{k})\cdot\Omega - I^{33}\sigma_1\big)\Omega + \Delta M\big((A_1\mathbf{k}\cdot v\big)v + eM^{11}\Omega\times v = t_1 A_1\mathbf{k}, \tag{4.22}$$

$$\big((\Delta I - e^2 M^{11})(A_2\mathbf{k})\cdot\Omega - I^{33}\sigma_2\big)\Omega + \Delta M\big((A_2\mathbf{k})\cdot v\big)v + eM^{11}\Omega\times v = t_2 A_2\mathbf{k}, \tag{4.23}$$

$$\Big(\mathbf{1} + \frac{\Delta M}{2M^{11}}(A_1\mathbf{k}\mathbf{k}^t A_1^t + A_2\mathbf{k}\mathbf{k}^t A_2^t)\Big)v = t_3\Omega + \frac{e}{2}\Omega\times(A_1\mathbf{k} + A_2\mathbf{k}), \tag{4.24}$$

such that no two solutions in the set can be obtained from each other by multiplying by an element of the symmetry group $SE(3)\times SO(2)^2$, and such that every solution can be obtained by multiplying one in the set by some element of $SE(3)\times SO(2)^2$. Then we have all relative equilibria up to the symmetry, with no replications.

This is accomplished by solving equations (4.22)–(4.24) on a subset of the set $\big\{(A_1, a_1, A_2, a_2, v, \Omega, \sigma_1, \sigma_2)\big\}$ which meets each orbit of the action

$$(A_1, a_1, A_2, a_2, \Omega, v) \mapsto (AA_1\exp(-\theta_1\mathbf{k}), Aa_1 + a, AA_2\exp(-\theta_2\mathbf{k}), Aa_2 + a, A\Omega, -(A\Omega)\times a + Av)$$

exactly once. To find such a set, think of standardizing configurations of the vehicles using re-orientation and translation. First, use translation to place the ball joint connecting the vehicles at the origin. That is, by choice of a, one may assume that

$$eA_1\mathbf{k} + a_1 = 0 = eA_2\mathbf{k} + a_2.$$

Then, using A, the first vehicle may be assumed to have the same orientation as its reference i.e. one may assume that $A_1 = \mathbf{1}$. The symmetry axis of the second vehicle then points in some arbitrary position. However, by choice of θ_1 and adjustment of A, it may be brought into the plane perpendicular to $\mathbf{j}$. By further choice of θ_2, we can then assume that $A_2 = \exp(\theta\mathbf{j})$ for some $\theta \in [0, \pi]$. In these configurations, θ is the angle at the joint between the two axes of symmetry of the vehicles. There is a further freedom when $\theta = 0$ or $\theta = \pi$, because for these configurations there is a further $SO(2)$ symmetry of rotation about the $\mathbf{k}$ axis. This can be used to arrange $\Omega^1 > 0$ and $\Omega^2 = 0$ as long as Ω is not parallel to $\mathbf{k}$. If Ω is parallel to $\mathbf{k}$ then one can arrange $v^1 > 0$ and $v^2 = 0$. The final case occurs where $\theta = 0$ or $\theta = \pi$ and both Ω and v are parallel to $\mathbf{k}$. To summarize, we will solve equations under the following seven cases:

$$A_1 = \mathbf{1}, \quad A_2 = \mathbf{1}, \quad \Omega = u^3\mathbf{k}, \quad v = v^3\mathbf{k}, \quad u^3, v^3 \in \mathbb{R}; \tag{5.1}$$

$$A_1 = \mathbf{1}, \quad A_2 = \exp(\pi\mathbf{j}^\wedge), \quad \Omega = u^3\mathbf{k}, \quad v = v^3\mathbf{k}, \quad u^3, v^3 \in \mathbb{R}; \tag{5.2}$$

$$A_1 = \mathbf{1}, \quad A_2 = \mathbf{1}, \quad \Omega = u^3\mathbf{k}, \quad v^1 > 0, \quad v^2 = 0, \quad u^3 \in \mathbb{R}; \tag{5.3}$$

$$A_1 = \mathbf{1}, \quad A_2 = \exp(\pi\mathbf{j}^\wedge), \quad \Omega = u^3\mathbf{k}, \quad v^1 > 0, \quad v^2 = 0, \quad u^3 \in \mathbb{R}; \tag{5.4}$$

$$A_1 = \mathbf{1}, \quad A_2 = \mathbf{1}, \quad \Omega^1 > 0, \quad \Omega_2 = 0; \tag{5.5}$$

$$A_1 = \mathbf{1}, \quad A_2 = \exp(\pi\mathbf{j}^\wedge), \quad \Omega^1 > 0, \quad \Omega_2 = 0; \tag{5.6}$$

$$A_1 = \mathbf{1}, \quad A_2 = \exp(\theta\mathbf{j}^\wedge), \quad 0 < \theta < \pi. \tag{5.7}$$

It is convenient to put

$$\Omega = \begin{bmatrix} u^1 \\ u^2 \\ u^3 \end{bmatrix}, \qquad v = \begin{bmatrix} v^1 \\ v^2 \\ v^3 \end{bmatrix}.$$

5.1 Preliminary to cases (5.1)–(5.6)

In all the cases (5.1)–(5.6), $A_1 = \mathbf{1}$ and

$$A_2 = \begin{bmatrix} \mu & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mu \end{bmatrix}$$

where $\mu = 1$ if $\theta = 0$ and $\mu = -1$ if $\theta = \pi$. Solving equation (4.24) for v gives

$$v^1 = \frac{t_3 u^1}{2M^{11}}, \quad v^2 = -\frac{(1+\mu)eu^1}{2}, \quad v^3 = \frac{t_3 u^3}{2M^{33}}. \tag{5.8}$$

Substituting (5.8) into (4.22) and (4.23) ought to yield four equations but two are identical, resulting in the three equations

$$I^{33}u^1\sigma_1 - \frac{\Delta M u^1 u^3 {t_3}^2}{4M^{33}M^{11}} - \frac{1}{2}\big(M^{11}e^2(\mu-1) + 2\Delta I\big)u^1u^3 = 0, \tag{5.9}$$

$$(\sigma_1 - \mu\sigma_2)I^{33}u^1 = 0, \tag{5.10}$$

$$\frac{e\Delta M u^1 u^3 t_3(\mu - 1)}{4M^{33}} = 0. \tag{5.11}$$

(5.8)–(5.11) are to be specialized to the cases (5.1)–(5.6) and then solved for

$$u^1, u^2, u^3, v^1, v^2, v^3, \sigma_1, \sigma_2.$$

Case (5.1)

Case (5.1) is

$$u^1 = u^2 = 0, \quad v^1 = v^2 = 0, \quad \mu = 1.$$

Substitution into (5.8)–(5.11) gives the single equation

$$v^3 = \frac{u^3 t_3}{2M^{33}}.$$

If $u^3 = 0$ then $v^3 = 0$ while if $u^3 \neq 0$ then v^3 can be any real number, since t_3 is any real number. This gives the solutions

$$A_1 = \mathbf{1}, \quad A_2 = \mathbf{1}, \quad \Omega = 0, \quad v = 0, \quad \sigma_1, \sigma_2 \in \mathbb{R},$$

and

$$A_1 = \mathbf{1}, \quad A_2 = \mathbf{1}, \quad \Omega = u^3\mathbf{k}, \quad v = v^3\mathbf{k}, \quad \sigma_1, \sigma_2 \in \mathbb{R}, \quad u^3 \neq 0, \quad v^3 \in \mathbb{R}.$$

Case (5.2)

Case (5.2) is

$$u^1 = u^2 = 0, \quad v^1 = v^2 = 0, \quad \mu = -1.$$

Substitution into (5.8)–(5.11) gives the single equation

$$v^3 = \frac{u^3 t_3}{2M^{33}}.$$

If $u^3 = 0$ then $v^3 = 0$ while if $u^3 \neq 0$ then v^3 can be any real number, since t_3 is any real number. This gives the solutions

$$A_1 = \mathbf{1}, \quad A_2 = \exp(\pi \mathbf{j}^\wedge), \quad \Omega = 0, \quad v = 0, \quad \sigma_1, \sigma_2 \in \mathbb{R}$$

and

$$A_1 = \mathbf{1}, \quad A_2 = \exp(\pi \mathbf{j}^\wedge), \quad \Omega = u^3 \mathbf{k}, \quad v = v^3 \mathbf{k}, \quad \sigma_1, \sigma_2 \in \mathbb{R}, \quad u^3 \neq 0, \quad v^3 \in \mathbb{R}.$$

Case (5.3)

Since $u^1 = 0$, we obtain from (5.8) $v_1 = 0$, so there are no solutions from this case since it assumes $v^1 > 0$.

Case (5.4)

Since $u^1 = 0$, we obtain from (5.8) $v_1 = 0$, so there are no solutions from this case since it assumes $v^1 > 0$.

Case (5.5)

In this case, $u^1 > 0$, $u^2 = 0$, and $\mu = 1$. Equation (5.11) is satisfied and (5.9), (5.10) become

$$u^1 \left(\Delta I u^3 - I^{33} \sigma_1 + \frac{\Delta M u^3 {t_3}^2}{4M^{11} M^{33}} \right) = 0, \qquad u^1 I^{33} (\sigma_1 - \sigma_2) = 0.$$

Since $u^1 > 0$, the second gives $\sigma_1 = \sigma_2$ and the remaining equation

$$I^{33} \sigma_1 - \left(\frac{\Delta M}{4M^{11} M^{33}} {t_3}^2 + \Delta I \right) u^3 = 0. \tag{5.12}$$

This is solved in the following cases, as follows: in each case,

$$v^1 = \frac{t_3 u^1}{2M^{11}}, \quad v^2 = -eu^1, \quad v^3 = \frac{t_3 u^3}{2M^{33}}.$$

1. If $\Delta M = 0$ and $\Delta I = 0$ then $\sigma_1 = 0$ from (5.12) and so $\sigma_1 = \sigma_2 = 0$. This gives the solution

$$\Delta M = 0, \quad \Delta I = 0, \quad \sigma_1 = 0, \quad \sigma_2 = 0, \quad t_3, u^1, u^3, \sigma_1, \sigma_2 \in \mathbb{R}, \quad u^1 > 0.$$

2. If $\Delta M = 0$ but $\Delta I \neq 0$, then $u^3 = I^{33}\sigma_1/\Delta I$, giving the solution

$$\Delta M = 0, \quad u^2 = 0, \quad u^3 = \frac{I^{33}\sigma_1}{\Delta I}, \quad \sigma_1 = \sigma_2, \quad u^1, t_3, \sigma_2 \in \mathbb{R}, \quad u^1 > 0.$$

3. If $\Delta M \neq 0$ then there are two subcases. If

$$t_3{}^2 = \frac{-4M^{33}M^{11}\Delta I}{\Delta M}$$

which is possible if and only if ΔM and ΔI have opposite sign, then $\sigma_1 = \sigma_2 = 0$, leading to the solution

$$u^2 = 0, \quad t_3 = \pm\sqrt{\frac{-4M^{11}M^{33}\Delta I}{\Delta M}}, \quad \sigma_1 = \sigma_2 = 0, \quad u^1, u^3 \in \mathbb{R}, \quad u^1 > 0.$$

If

$$t_3{}^2 \neq \frac{-4M^{33}M^{11}\Delta I}{\Delta M}$$

then there is the solution

$$u^3 = \frac{4M^{11}M^{33}I^{33}\sigma_1}{4M^{11}M^{33}\Delta I + \Delta M t_3{}^2}, \quad \sigma_1 = \sigma_2, \quad u^1, \sigma_2, t_3 \in \mathbb{R}, \quad u^1 > 0.$$

Case (5.6)

In this case $u^1 > 0$, $u^2 = 0$, and $\mu = -1$. Equation (5.10) gives $\sigma_1 = \sigma_2$ and (5.11) becomes

$$\frac{u^1 u^3 t_3 e \Delta M}{2M^{33}} = 0$$

from which one has the following three cases. In each case,

$$v^1 = \frac{t_3 u^1}{2M^{11}}, \quad v^2 = 0, \quad v^3 = \frac{t_3 u^3}{2M^{33}}.$$

1. If $\Delta M = 0$ or $t_3 = 0$ then (5.9) gives

$$I^{33}\sigma_1 + (M^{11}e^2 - \Delta I)u^3 = 0.$$

   If $\Delta I = M^{11}e^2$ then this implies $\sigma_1 = 0$, leading to the solutions

$$\Delta M = 0, \quad \Delta I = M^{11}e^2, \quad u^2 = 0, \quad \sigma_1 = 0, \quad \sigma_2 = 0, \quad u^1, u^3, t_3 \in \mathbb{R}, \quad u^1 > 0,$$
$$t_3 = 0, \quad \Delta I = M^{11}e^2, \quad u^2 = 0, \quad \sigma_1 = 0, \quad \sigma_2 = 0, \quad u^1, u^3 \in \mathbb{R}, \quad u^1 > 0.$$

   If $\Delta I \neq M^{11}e^2$ then there are the solutions

$$\Delta M = 0, \quad u^2 = 0, \quad u^3 = \frac{I^{33}\sigma_1}{\Delta I - M^{11}e^2}, \quad \sigma_1 = -\sigma_2, \quad \sigma_2, u^1, t_3 \in \mathbb{R}, \quad u^1 > 0,$$
$$t_3 = 0, \quad u^2 = 0, \quad u^3 = \frac{I^{33}\sigma_1}{\Delta I - M^{11}e^2}, \quad \sigma_1 = -\sigma_2, \quad \sigma_2, u^1 \in \mathbb{R}, \quad u^1 > 0.$$

2. If $u_3 = 0$ then (5.11) is satisfied and (5.9) is $u^1 I^{33}\sigma_1 = 0$ i.e. $\sigma_1 = 0$. This gives the solution

$$u^2 = 0, \quad u^3 = 0, \quad \sigma_1 = 0, \quad \sigma_2 = 0, \quad u^1, t_3 \in \mathbb{R}, \quad u^1 > 0.$$

Case (5.7)

Case (5.7), subcase $\Omega \not\parallel A_1\mathbf{k}, \Omega \not\parallel A_2\mathbf{k}$

In this case Ω is not a multiple of either $A_1\mathbf{k}$ not $A_2\mathbf{k}$, so, in particular, $\Omega \neq 0$. The variable t_1 in equation (4.22) is eliminated by noting that (4.22) is equivalent to the vanishing of its dot product with $A_1\mathbf{i}$ and $A_1\mathbf{j}$ i.e. it is equivalent to

$$\begin{aligned}&\big((\Delta I - e^2M^{11})(A_1\mathbf{k})\cdot\Omega - I^{33}\sigma_1\big)\Omega\cdot(A_1\mathbf{i})\\&\qquad + \Delta M\big((A_1\mathbf{k})\cdot v\big)\big(v\cdot(A_1\mathbf{i})\big) + eM^{11}(\Omega\times v)\cdot(A_1\mathbf{i}) = 0, \qquad (5.13)\\&\big((\Delta I - e^2M^{11})(A_1\mathbf{k})\cdot\Omega - I^{33}\sigma_1\big)\Omega\cdot(A_1\mathbf{j})\\&\qquad + \Delta M\big((A_1\mathbf{k})\cdot v\big)\big(v\cdot(A_1\mathbf{j})\big) + eM^{11}(\Omega\times v)\cdot(A_1\mathbf{j}) = 0. \qquad (5.14)\end{aligned}$$

From (5.13) and (5.14), σ_1 is eliminated by $\Omega\cdot A_1\mathbf{j}$ times the (5.13) minus $\Omega\cdot A_1\mathbf{i}$ times (5.14), giving

$$\Delta M\big((A_1\mathbf{k})\cdot v\big)\big((v\cdot(A_1\mathbf{i}))(\Omega\cdot(A_1\mathbf{j})) - (v\cdot(A_1\mathbf{j}))(\Omega\cdot(A_1\mathbf{i}))\big)$$

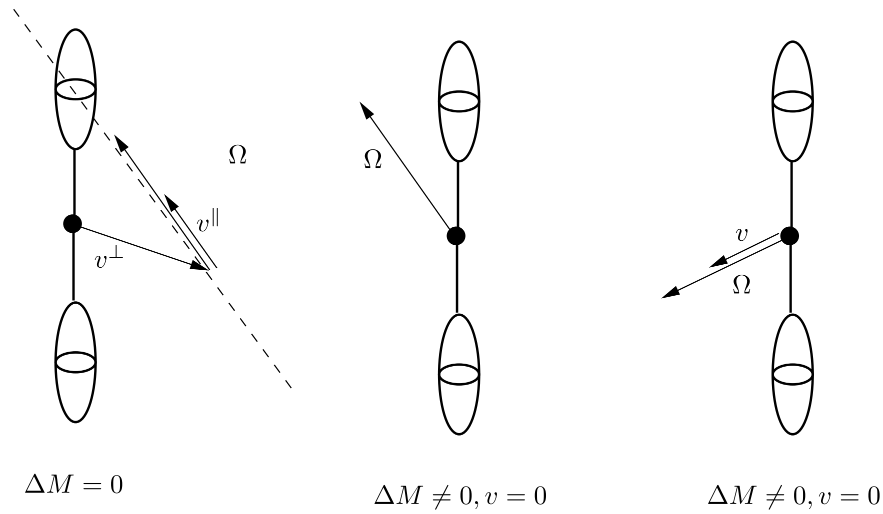

Figure 5.1: Illustrating the solutions from Case (5.7). The vehicles are in the opposed configuration and are spinning together if $\sigma_1 = -\sigma_2 \neq 0$. On the left, $v \neq 0$ and the vehicles move in a screw motion along the axis indicated by the dashed line. On the middle, $v = 0$, and there is no translation. On the right, v and Ω are both parallel to $\mathbf{i}$ and the vehicles rotate and uniformly translate about the axis $\mathbb{R}\mathbf{i}$.

$$
\begin{aligned}
&\quad + eM^{11}\big(((\Omega \times v) \cdot (A_1\mathbf{i}))((A_1\mathbf{j}) \cdot \Omega) - ((\Omega \times v) \cdot (A_1\mathbf{j}))((A_1\mathbf{i}) \cdot \Omega)\big) \\
&= \Delta M\big((A_1\mathbf{k}) \cdot v\big)\big((v \times \Omega) \cdot ((A_1\mathbf{i}) \times (A_1\mathbf{j}))\big) \\
&\quad + eM^{11}\big(((\Omega \times v) \times \Omega) \cdot ((A_1\mathbf{i}) \times (A_1\mathbf{j}))\big) \\
&= \left(\Delta M\big((A_1\mathbf{k}) \cdot v\big)(v \times \Omega) + eM^{11}\big((\Omega \times v) \times \Omega\big)\right) \cdot A_1\mathbf{k} \\
&= 0.
\end{aligned}
$$

Similarly t_2 and σ_2 are eliminated from (4.23). The result is that (4.22) and (4.23) imply

$$\left(\Delta M\big((A_1\mathbf{k}) \cdot v\big)(v \times \Omega) + eM^{11}\big((\Omega \times v) \times \Omega\big)\right) \cdot A_1\mathbf{k} = 0, \tag{5.15}$$

$$\left(\Delta M\big((A_2\mathbf{k}) \cdot v\big)(v \times \Omega) + eM^{11}\big((\Omega \times v) \times \Omega\big)\right) \cdot A_2\mathbf{k} = 0. \tag{5.16}$$

Under the assumption $\Omega \not\parallel A_1\mathbf{k}$, (5.15) implies that (4.22) is solved by exactly one σ_1, so the vector equation (4.22) can be replaced with (5.15) given that (4.22) is used to determine σ_1. Similarly, equation (4.23) can be replaced with (5.16).

The sum of equations (5.15) and (5.16) is

$$\begin{aligned}\Delta M\big((A_1\mathbf{k}) \cdot v\big)\big((v \times \Omega) \cdot A_1\mathbf{k}\big) + \Delta M\big((A_2\mathbf{k}) \cdot v\big)\big((v \times \Omega) \cdot A_2\mathbf{k}\big) \\ + eM^{11}(\Omega \times v) \cdot \big(\Omega \times (A_1\mathbf{k} + A_2\mathbf{k})\big) = 0.\end{aligned} \tag{5.17}$$

If we assume (4.24), then

$$\frac{e}{2}\Omega \times (A_1\mathbf{k} + A_2\mathbf{k}) = -t_3\Omega + \Big(\mathbf{1} + \frac{\Delta M}{2M^{11}}(A_1\mathbf{k}\mathbf{k}^t A_1^t + A_2\mathbf{k}\mathbf{k}^t A_2^t)\Big)v$$

and substituting this into (5.17) gives $0 = 0$, so (5.15) and (5.16) are redundant under (4.24). Discarding (5.16), and using $A_1 = \mathbf{1}$, all solutions from this subcase are thus given by solutions of (4.24) and

$$\big(\Delta M(\mathbf{k} \cdot v)\mathbf{k} + eM^{11}\mathbf{k} \times \Omega\big) \cdot (v \times \Omega) = 0. \tag{5.18}$$

Into (5.18) one is to substitute v from (4.24). The result is an algebraically complicated single equation which is at most quadratic in t_3. Because of this complexity, we only give a procedure that finds all solutions in this subcase:

1. Pick any $0 < \theta < \pi$ and any Ω with $\Omega \not\parallel A_1 = k$ and $\Omega \not\parallel A_2k$; set $A_1 = \mathbf{1}$ and $A_2 = \exp(\theta\mathbf{j}^\wedge)$;

2. Calculate
$$C(\theta) = \Big(\mathbf{1} + \frac{\Delta M}{2M^{11}}(A_1\mathbf{k}\mathbf{k}^t A_1^t + A_2\mathbf{k}\mathbf{k}^t A_2^t)\Big);$$
and calculate
$$v = C(\theta)^{-1}\left((t_3\Omega + \frac{e}{2}\Omega \times (A_1\mathbf{k} + A_2\mathbf{k})\right), \tag{5.19}$$
giving v as a linear function in t_3;

3. substitute Ω and v into (5.18) and solve for t_3; if there are no solutions then there are no relative equilibria the choice of θ, and Ω;

4. substitute t_3 into (5.19) and calculate v;

5. substitute Ω, v, and t_3 into (4.22) and (4.23) and calculate σ_1 and σ_2.

Case (5.7), subcase $\Omega \parallel A_1\mathbf{k}, \Omega \not\parallel A_2\mathbf{k}$

Since $\Omega \parallel A_1\mathbf{k}$ we have $\Omega = u^3\mathbf{k}$ i.e. $u^1 = 0$ and $u^2 = 0$. Substituting this into (4.22) and dotting with $\mathbf{i}$ and $\mathbf{j}$ gives the two equations

$$\Delta M v^1 v^3 - eM^{11} u^3 v^2 = 0, \quad \Delta M v^2 v^3 + eM^{11} u^3 v^1 = 0,$$

respectively. v^1 times the second minus v^2 times the first gives

$$eM^{11} u^3 \big((v^1)^2 + (v^2)^2\big) = 0$$

which gives the cases $v^1 = v^2 = 0$ or $u^3 = 0$. t_3 in (4.24) absorbs any mutiple of $A_1\mathbf{k}$, so (4.24) is equivalent to

$$\Big(\mathbf{1} + \frac{\Delta M}{2M^{11}}(A_2\mathbf{k}\mathbf{k}^t A_2^t)\Big)v = t_3\Omega + \frac{e}{2}\Omega \times (A_1\mathbf{k} + A_2\mathbf{k}). \tag{5.20}$$

Putting $\Omega = u^3\mathbf{k}$ into (5.20) and absorbing u^3 into t_3 gives

$$\Big(\mathbf{1} + \frac{\Delta M}{2M^{11}}(A_2\mathbf{k}\mathbf{k}^t A_2^t)\Big)v = t_3\mathbf{k} + u^3\frac{e}{2}\mathbf{k} \times (A_2\mathbf{k}). \tag{5.21}$$

Dotting both sides of (5.21) $\mathbf{k} \times (A_2\mathbf{k}) = -\sin\theta\mathbf{j}$ gives

$$-v^2 \sin\theta = \frac{eu^3 \sin^2\theta}{2}.$$

Thus the case $v^1 = v^2 = 0$ leads to $u^3 = 0$ as well, so $\Omega = 0$ and hence $v = 0$, and the solutions from this subcase are

$$A_1 = \mathbf{1}, \quad A_2 = \exp(\theta\mathbf{j}^\wedge), \quad \Omega = 0, \quad v = 0, \quad \sigma_1, \sigma_2 \in \mathbb{R}, \quad 0 < \theta < \pi. \tag{5.22}$$

5.2 Existence of relative equilibria of type $\mu_e^a = 0$, $\xi_e^r = 0$, $\mu_e^r \neq 0$

The notational correspondence with [13] is

$$\xi_e^r \leftrightarrow \Omega, \quad \xi_e^a \leftrightarrow v, \quad \mu_e^r \leftrightarrow \Omega, \quad \mu_e^a \leftrightarrow p$$

(5.22) have $\Omega = 0$, corresponding to $\xi_e^r = 0$. Substituting (5.22) into (4.14) and (4.15) gives

$$\Omega_i = A_i^t\Omega - \sigma_i\mathbf{k} = -\sigma_i\mathbf{k},$$

$$v_i = A_i^t(\Omega \times a_i + v) = 0.$$

This into the momentum (4.17) gives $p = 0$, corresponding to $\mu_e^a = 0$, and

$$\mu = A_1 I_1(-\sigma_1 \mathbf{k}) + A_2 I_2(-\sigma_2 \mathbf{k}) = -I^{33}\big(\sigma_1 \mathbf{k} + \sigma_2 \exp(\theta \mathbf{j}^\wedge)\mathbf{k}\big).$$

Thus, for any $\theta \neq 0$ or $\theta \neq \pi$, μ is any linear combinations (σ_1, σ_2 are arbitrary) of two linearly independent vectors, so one certainly can arrange $\mu \neq 0$. Since this corresponds to $\mu_e^r \neq 0$, there are relative equilibria of the required type.

5.3 Existence of relative equilibria of type $\mu_e^a = 0$, $\xi_e^r \neq 0$, $\mu_e^r = 0$

$\mu_e^a = 0$ and $\mu_e^r = 0$ correspond to $\mu = 0$ and $p = 0$, i.e.

$$\sum_{i=1}^{2} A_i I \Omega_i + a_i \times (A_i M v_i) = 0, \quad \sum_{i=1}^{2} A_i M v_i = 0.$$

Substituting

$$A_1 = \mathbf{1}, \quad A_2 = \exp(\theta \mathbf{j}^\wedge), \quad \Omega_i = A_1^t \Omega - \sigma_i \mathbf{k}, \quad v_i = A_i^t(\Omega \times a_1 + v)$$

yields linear equations in Ω and v of the form

$$D \begin{bmatrix} \Omega \\ v \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -I^{33}\sigma_2 \sin\theta \\ 0 \\ -I^{33}(\sigma_1 + \sigma_2 \cos\theta) \end{bmatrix} = 0 \tag{5.23}$$

where D is a 6×6 matrix and $\det D = 0$ if and only if $\sec\theta$ is equal to one of

$$\frac{M^{11} - M^{33}}{M^{11} + M^{33}}, \quad \frac{I^{11} - I^{33}}{I^{11} + I^{33}}, \quad \frac{I^{11} + e^2 M^{11} - I^{33}}{I^{11} + e^2 M^{11} + I^{33}}, \quad \frac{M^{33}(I^{11} + e^2 M^{11}) - I^{11} M^{11}}{M^{33}(I^{11} + e^2 M^{11}) + I^{11} M^{11}}.$$

This is not possible since every number in the list is smaller than 1. So (5.23) can be solved for Ω and v, the result being linear in σ_1 and σ_2. Substituting the result into equations (4.22)–(4.24) and putting $t_3 = 0$ yields a single quadratic equation in σ_1 and σ_2. The discriminant of that quadratic is

$$\frac{4I^{11}(I^{33})^4(I^{11}+e^2M^{11})\sin^2\theta}{(I^{11}+I^{33}+(I^{11}-I^{33})\cos\theta)^2\big(I^{11}+e^2M^{11}+I^{33}-(I^{11}+e^2M^{11}-I^{33})\cos\theta\big)^2}.$$

Since this is positive, there are solutions σ_1 and σ_2 to the quadratic, and hence there are relative equilbria of the required type.

Bibliography

[1] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. Addision-Wesley, second edition, 1978.

[2] R. Abraham, J. E. Marsden, and T. S. Ratiu. *Manifolds, tensor analysis, and applications*. Springer-Verlag, second edition, 1988.

[3] P. Holmes, J. Jenkins, and N. E. Leonard. Dynamics of the Kirchoff equations I: Coincident centers of gravity and bouyancy. *Physica D*, 118:311–342, 1998.

[4] N. E. Leonard. Stability of a bottom-heavy underwater vehicle. *Automatica—J. IFAC*, 33:331–346, 1997.

[5] N. E. Leonard and J. E. Marsden. Stability and drift of underwater vehicle dynamics: mechanical systems with rigid motion symmetry. *Physica D*, 105:130–162, 1997.

[6] J. E. Marsden, J. C. Simo, D. Lewis, and T. A. Posbergh. Block diagonalization and the energy-momentum method. In J. E. Marsden, P. S. Krishnaprasad, and J. C. Simo, editors, *Dynamics and Control of Multibody Systems, Contemporary Mathematics 97*, pages 297–313. AMS, 1989.

[7] G. W. Patrick. *Two axially symmetric coupled rigid bodies: relative equilibria, stability, bifurcations, and a momentum preserving symplectic integrator*. PhD thesis, University of California at Berkeley, 1991.

[8] G. W. Patrick. Relative equilibria in Hamiltonian systems: The dynamic interpretation of nonlinear stability on the reduced phase space. *J. Geom. Phys.*, 9:111–119, 1992.

[9] G. W. Patrick. Relative equilibria of Hamiltonian systems with symmetry: linearization, smoothness, and drift. *J. Nonlinear Sci.*, 5:373–418, 1995.

[10] G. W. Patrick. Dynamics near stable relative equilibria at non-generic momenta: a numerical investigation. *Fields Inst. Comm.*, 8:127–142, 1996.

[11] G. W. Patrick. Dynamics near relative equilibria: Nongeneric momenta at a 1:1 group–reduced resonance. *Math. Z.*, 232:747–788, 1999.

[12] G. W. Patrick. Stability by KAM confinement of certain wild, nongeneric relative equilibria of underwater vehicles with coincident centers of mass and bouyancy. *SIAM J. Appl. Dyn. Syst.*, 2:36–52, 2003.

[13] G. W. Patrick, R. M. Roberts, and C. Wulff. Stability of Poisson equilibria and Hamiltonian relative equilibria by energy methods. *Arch. Rational Mech. Anal.*, 174:301–344, 2004.

Index