STABILITY OF GENERIC EQUILIBRIA OF THE 2N DIMENSIONAL FREE RIGID BODY USING THE ENERGY-CASIMIR METHOD

by

Adam Howard Spiegler

A Dissertation Submitted to the Faculty of the
DEPARTMENT OF MATHEMATICS
In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY
In the Graduate College
THE UNIVERSITY OF ARIZONA

2006
As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Adam Howard Spiegler, entitled Stability of Generic Equilibria of the $2n$ Dimensional Free Rigid Body Using the Energy-Casimir Method and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Date: 28 April 2006
Hermann Flaschka

Date: 28 April 2006
Nicholas Ercolani

Date: 28 April 2006
Doug Pickrell

Date: 28 April 2006
Philip Foth

Date: 28 April 2006
Vladimir Zakharov

Final approval and acceptance of this dissertation is contingent upon the candidate’s submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: 28 April 2006
Dissertation Director: Hermann Flaschka
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: 

______________________________
Adam Howard Spiegler
ACKNOWLEDGEMENTS

There are many people that I would like to thank for making this dissertation possible. First, I owe an infinite amount of gratitude to my advisor Hermann Flaschka for his guidance, encouragement, and patience over the past couple of years. His willingness to make time and check that I stay on course has been extremely generous and helpful. I am also indebted for his thorough proofreading of this dissertation. Thank you for making this entire experience rewarding and enjoyable.

I thank Tudor Ratiu for originally proposing this problem to us and introducing me to the field of geometric mechanics. The work has been extremely engaging and fruitful. I am grateful for his continued interest in our work, for his guidance, and for his time reviewing this dissertation.

I thank my final defense committee: Nick Ercolani, Vladimir Zakharov, Doug Pickrell, and Philip Foth. Their time checking my work and their constructive feedback are greatly appreciated. The entire faculty of the Mathematics Department at the University of Arizona has been very helpful. In particular, I would like to thank Michael Otto for our many mathematical discussions. I would like to thank Tina Deemer, David Gay, Bruce Bayly, and Deborah Hughes-Hallett for their invaluable guidance, supervision, and interest in my teaching.

My family has been a constant source of support and motivation. I have been inspired by the tremendous strength of my grandparents. Without their sacrifices and determination, I would not be in this fortunate position today. My parents, Matt, Jana, Justin, Alice, and Arthur for their unconditional support and love. Without the foundation of my family, none of this would be possible.

There are many classmates and friends that I thank for their support. I thank Emily Landeen for her love, support, patience, and sacrifice. Words cannot express how much she has helped me the past couple of years. She is my shelter from the storm. I thank Panagiota Konstantinou for her friendship, compassion, and support.

I would like to thank my officemates Eric Roberts, Nakul Chitnis, Luis Garcia Naranjo, Andy Linfoot, Yijun Shao, and Alan Von Herrmann. We have had many spirited discussions about mathematics and more importantly life beyond the chalkboards and walls of our office. They have certainly made the experience very special. I will miss our daily discussions enormously.

Thank you Davide Dattilo, Antonio Colangelo, Mariagrazia Mecoli, Oliverio Alvarez, the Green Team, Derek Habermas, Lisa Berger, Steve Foster, Daniel Reich, and Ben Levitt for helping make my time in Tucson enjoyable. Thank you to Cyril Derzie and Ofer Wolberger for helping me travel freely through this world and explore the beauty of life. Thank you to Dan D’Orazio, Seth Barish, Brent Fisher, and Matt Lefever for your long and devout friendships.
DEDICATION

To my parents Neil and Peggy Spiegler. Thank you for believing in me and encouraging me to persue my dreams.
# TABLE OF CONTENTS

LIST OF FIGURES ......................................................... 8

ABSTRACT ................................................................. 9

CHAPTER 1 INTRODUCTION ............................................... 10
  1.1 The Geometry of the Rigid Body ................................. 13
    1.1.1 Body and Space Coordinates .............................. 14
    1.1.2 Hamiltonian Vector Fields on the Cotangent Bundle ...... 17
    1.1.3 Coadjoint Orbit Reduction ................................. 18
    1.1.4 General Symplectic Reduction ............................. 20
    1.1.5 The Inertia Map ............................................ 21
    1.1.6 Identifying Coadjoint and Adjoint Orbits ................. 24
  1.2 Deriving Euler’s Equation ...................................... 26
    1.2.1 Identifying Angular Velocity and Angular Momentum .... 26
    1.2.2 The Energy Metric ........................................ 27
    1.2.3 The Inertia Map of the Rigid Body ....................... 29
    1.2.4 The Equations of Motion .................................. 32
    1.2.5 Example: The 3 Dimensional Case ......................... 33
  1.3 Integrability .................................................... 35

CHAPTER 2 REDUCED DYNAMICS: THE ENERGY-CASIMIR METHOD AND CLASSIFICATION OF GENERIC EQUILIBRIA .......... 40
  2.1 Structure of Non-Degenerate Adjoint Orbits .................... 41
  2.2 Relative Equilibria ............................................. 42
  2.3 The Energy-Casimir Method .................................... 43
  2.4 Generic equilibria ............................................. 46

CHAPTER 3 THE ENERGY-CASIMIR METHOD AND STABILITY OF GENERIC EQUILIBRIA OF THE 2N DIMENSIONAL RIGID BODY .... 58
  3.1 A General Family of Casimirs ................................ 60
  3.2 The First Variation ............................................ 62
  3.3 The Second Variation ......................................... 68
    3.3.1 A Basis for Setting Up a Matrix Representation for $d^2 H_\phi(\Omega_e)$ 69
    3.3.2 Constructing the Matrix of $d^2 H$ ...................... 74
    3.3.3 The Matrix of First Partial Derivatives ................. 76
    3.3.4 Constructing the Matrix of Second Partialials .......... 88
TABLE OF CONTENTS – Continued

3.4 Definiteness of the Second Variation ........................................ 92
  3.4.1 Definiteness of the $2 \times 2$ Blocks ............................... 93
  3.4.2 Definiteness of the $n \times n$ Block ............................... 105

APPENDIX A HISTORYCAL BACKGROUND OF THE RIGID BODY . . . 110

APPENDIX B THE CALCULATION OF THE PRODUCT
\[ d\tilde{F}_{\Omega_{g}} \cdot Hess(\phi) \cdot d\tilde{F}_{\Omega_{g}} \] IN THEOREM 3.3.8 ................................. 115

REFERENCES .............................................................................. 117
# LIST OF FIGURES

1.1 Body and space coordinates for a general Lie group $G$. The inertia map $\tilde{\mathcal{J}}$ connects the Ad-invariant induced pairing of $\mathfrak{g}$ and $\mathfrak{g}^*$ denoted $(\cdot, \cdot)$ (for example the Killing form) with the left invariant energy metric denoted $(\cdot, \cdot)$. Since the energy metric is left invariant, the inertia map $\tilde{\mathcal{J}}$ commutes with the left trivializations of $TG$ and $T^*G$, but not necessarily with the right trivializations. The map $\tilde{\mathcal{J}}_g$ denotes the left translation of the inertia map $\tilde{\mathcal{J}}$ from the identity to the tangent and cotangent spaces over an arbitrary $g \in G$. 

1.2 The trajectories of the three dimensional rigid body on an adjoint orbit.

2.1 An illustration of Arnold’s energy-Casimir method. Flows in the intersections of the symplectic leaves and level sets of modified Hamiltonian $\tilde{H} + C$ lie inside ellipsoid-like hypersurfaces, and remain in a neighborhood near the equilibrium $\mu_e$. [4]

A.1 An illustration of Poinsot’s construction
ABSTRACT

The rigid body has been one of the most noteworthy applications of Newtonian mechanics. Applying the principles of classical mechanics to the rigid body is by no means routine. The equations of motion, though discovered two hundred and fifty years ago by Euler, have remained quite elusive since their introduction. Understanding the rigid body has required the applications of concepts from integrable systems, algebraic geometry, Lie groups, representation theory, and symplectic geometry to name a few. Moreover, several important developments in these fields have in fact originated with the study of the rigid body and subsequently have grown into general theories with much wider applications.

In this work, we study the stability of equilibria of non-degenerate orbits of the generalized rigid body. The energy-Casimir method introduced by V.I. Arnold in 1966 allows us to prove stability of certain non-degenerate equilibria of systems on Lie groups. Applied to the three dimensional rigid body, it recovers the classical Euler stability theorem \[12\]: rotations around the longest and shortest principal moments of inertia are stable equilibria. This method has not been applied to the analysis of rigid body dynamics beyond dimension \(n = 3\). Furthermore, no conditions for the stability of equilibria are known at all beyond \(n = 4\), in which case the conditions are not of the elegant longest/shortest type \[10\].

Utilizing the rich geometric structures of the symmetry group \(G = SO(2n)\), we obtain stability results for generic equilibria of the even dimensional free rigid body. After obtaining a general expression for the generic equilibria, we apply the energy-Casimir method and find that indeed the classical longest/shortest conditions on the entries of the inertia matrix are sufficient to prove stability of generic equilibria for the generalized rigid body in even dimensions.
CHAPTER 1

INTRODUCTION

The study of the classical rigid body rotating freely around a fixed point in three-dimensional space has played an illustrious role in the advancement of classical mechanics. The legacy of the rigid body has unfolded simultaneously with the growth of mechanics. The equations of motion have led to some remarkable discoveries and applications both inside and outside of mathematics. Geometric, analytic, and algebraic techniques all have played important roles in the study of the dynamics of the rigid body, and interesting connections in these fields have evolved from their applications to the rigid body. Moreover, theories which have emerged from the study of the rigid body have been generalized to useful applications in areas such as celestial, quantum, and fluid mechanics as well as robotics, control theory, plasma physics, and areas of biology and chemistry.

The rigid body is the root of many interesting problems which are obtained by generalizing or considering special cases of the rigid body. In addition to generalizing the free rigid body to higher dimensions, the rigid body has been modified to include external forces, constraints, and various symmetries in the mass distribution. These subsequent problems have in turn significantly influenced the development of areas such as nonholonomic dynamics.

Due to the symmetry of the motion, the configuration space of the free $n$-dimensional rigid body is the Lie group $G = SO(n)$, the group of orthogonal rotations in $\mathbb{R}^n$. The phase space is thus the cotangent bundle $T^*G$, which has a canonical symplectic structure. As a result of the symmetry of the system (the energy is left invariant), we can pull the dynamics of the system to the fiber over the identity of $T^*G$, where we can further reduce the dynamics to the coadjoint orbits of $\mathfrak{g}^*$. The
rigid body is thus a key example of symplectic reduction which is referred to as coadjoint orbit reduction. In addition, the rigid body is a very special case of geodesic motion on a Lie group which is endowed with a one sided-invariant metric. For instance, the classical Euler equation governing the flow of an ideal (incompressible and inviscid) homogeneous fluid in a domain can be viewed as a generalization of the rigid body equation to the group of volume-preserving diffeomorphisms. Arnold shows the equations of such systems in hydrodynamics may be viewed as the geodesic sprays of a right invariant metric on the group of diffeomorphisms.

The equations of motion for the $n$ dimensional free rigid body can be written as a Lax pair $\dot{M} = [M, \Omega]$ of the angular momentum and velocities in the body frame, a generalization of Euler’s equation to higher dimensional space. Therefore, the equations are special since the eigenvalues of $M$ are constant, and symmetric functions of the eigenvalues will likewise be constant. The integrability of the three dimensional rigid body follows immediately since after the reduction process, it is a two dimensional Hamiltonian system on the coadjoint orbits. It is not clear in higher dimensions that we can generate enough integrals of the motion which will be in involution; yet the $n$ dimensional rigid body is integrable as well. The Lax pair equations of rigid body are a particular case a larger collection of integrable systems whose equations are of the form of the Dubrovin equation. This in turn leads to a beautiful application of algebraic geometry. The solutions can be found by integrating Abelian functions, and we can find an explicit (but extremely complicated) solution in terms of ratios of Riemann theta functions; however, the role of integrability does not play a role in our treatment of the rigid body equations.

In this work we study the stability of equilibria on non-degenerate orbits of the generalized rigid body. In three dimensions one has the classical Euler’s stability theorem which can be physically observed: spin a textbook around its three (distinct) axes and notice rotations around the longest and shortest axes will continue to rotate around those axes. Thus it is well established that rotations around the
longest and shortest principal moments of inertia are stable equilibria. Conditions for the four dimensional rigid body have been determined; however, the method Fehér and Marshall use to obtain these conditions is particular to the dimension $n = 4$ [10]. Moreover, their conditions are not of the elegant longest/shortest axis type.

The question of stability in $n$ dimensions is of mathematical interest because the answer should involve the inherent structure of the Lie group $SO(n)$. It is of physical interest because of the similarity between the Euler equation of $SO(n)$ and the Euler equation analogue of fluid dynamics. In the 1980’s and 1990’s, Ratiu and others extended and developed the so-called energy-Casimir method introduced by V.I. Arnold in 1966. This technique allows us to prove stability of certain non-degenerate equilibria of systems with symmetry. Applied to the three dimensional rigid body, it recovers the classical conditions. In our current work, we will utilize the geometric structures of the symmetry group $G = SO(2n)$ to obtain stability results for the free rigid body rotating in even dimensional space. The conditions are indeed the long/short axis type.

In this chapter, we begin in Section 1.1 by providing a quick glimpse into the various layers of geometry which allow us to reduce general systems with symmetry, whose configuration space is the symmetry group itself. The process is a special case of Marsden-Weinstein symplectic reduction [22], namely coadjoint orbit reduction applies to the case when the symplectic manifold is $T^* G$, and $H$ is a $G$-invariant Hamiltonian. Next, in Section 1.2 we apply these geometric tools to construct the equations of motion for the $n$ dimensional rigid body on the Lie algebra $\mathfrak{so}(n)$. The generalized Euler equation for the rigid body is derived in Theorem 1.2.2. Next we look at the special case of the three dimensional rigid body, and recover the classical Euler equation (1.39) in $\mathbb{R}^3$. Finally, though it will not play a role in our analysis, in Section 1.3 we end this chapter discussing the integrability of the rigid body, which is a special case of the Dubrovin equation, restricted from $\mathfrak{sl}(n)$ to $\mathfrak{so}(n)$. For a brief
historical review of the rigid body problem, we direct the reader to Appendix A.

1.1 The Geometry of the Rigid Body

The generalized rigid body problem deals with the study of a rigid body in $\mathbb{R}^n$ which rotates freely around the center of its mass. The configuration space is the space of orthogonal rotations in $\mathbb{R}^n$, the Lie group $G = SO(n)$. For mechanical systems in general, we can construct the equations of motion using the Lagrangian formulation and the associated Euler-Lagrange equations on $TG$. Or we can introduce a change of coordinates, the Legendre transform, and construct the corresponding Hamiltonian system on the phase space $T^*G$. Depending on the particular nature of the mechanical system, it may be more advantageous to setup the Lagrangian or the Hamiltonian. In the case of the rigid body, we will see the distinction between Lagrangian and Hamiltonian is a bit blurry after identifying $\mathfrak{g}$ with $\mathfrak{g}^*$, and we draw on both to ultimately construct the equations of motion.

In this section we provide a very brief overview of a particular case of symplectic reduction, coadjoint orbit reduction, which is useful for simplifying the dynamics of systems on Lie groups. In the case of the rigid body, $G = SO(n)$ and $H$ is the left invariant quadratic energy function; however, the reduction of $T^*G$ holds for any Lie group $G$ and any one-sided invariant smooth function $H : T^*G \to \mathbb{R}$. First we present body and space coordinates abstractly for arbitrary $TG$ and $T^*G$. Due to the left invariance of the energy functions of the rigid body, we then use body coordinates in $T^*G$ to translate the Hamiltonian system to $\mathfrak{g}^*$. Finally, in the special case when the corresponding Lie algebra $\mathfrak{g}$ of $G$ is semi-simple, we can identify $\mathfrak{g}$ with $\mathfrak{g}^*$. Using the inertia map we then construct the reduced Hamiltonian system on the adjoint orbits of $\mathfrak{g}$. The inertia essentially acts as the Legendre transform which allows us to pass the dynamics between the Lagrangian formulation in $\mathfrak{g}$ and the Hamiltonian formulation in $\mathfrak{g}^*$.
1.1.1 Body and Space Coordinates

Let $G$ be a finite dimensional Lie group, and $\mathfrak{g}$ its Lie algebra. Consider the tangent bundle $TG$ which can be identified with $G \times \mathfrak{g}$ in the following fashion. Given a point $v_h \in TG$, we can translate back to the fiber over the identity by the action of group multiplication on the left, which is denoted

$$L_g : G \to G : h \mapsto gh.$$ (1.1)

The derivative of this action, $T_h L_g : T_h G \to T_{gh} G$, induces a map on the corresponding tangent fibers. In particular if we set $g = h^{-1}$, then $T_h L_{h^{-1}}$ will map into $T_e G \cong \mathfrak{g}$.

**Definition 1.1.1** The identification of the tangent bundle with the Lie algebra induced by the push forward of left translation, $T_g L_{g^{-1}}$, is called the **left trivialization** of the tangent bundle and is denoted by

$$\lambda : TG \to G \times \mathfrak{g} : v_g \mapsto (g, T_g L_{g^{-1}}(v_g))$$ (1.2)

In the same manner, we could choose to identify $TG$ with $G \times \mathfrak{g}$ by multiplication from the right $R_g$. Again, the derivative of right multiplication induces a map on the corresponding tangent spaces.

**Definition 1.1.2** The identification of the tangent bundle with the Lie algebra induced by the push forward of right translation, $T_g R_{g^{-1}}$, is called the **right trivialization** and denoted

$$\rho : TG \to G \times \mathfrak{g} : v_g \mapsto (g, T_g R_{g^{-1}}(v_g))$$ (1.3)
The transition from body to space coordinates is given by the map $\rho \circ \lambda^{-1}$,

$$(\rho \circ \lambda^{-1}) : G \times \mathfrak{g} \to G \times \mathfrak{g} : (g, \xi) \mapsto (g, T_g R_{g^{-1}} \circ T_g L_g(\xi)) = (g, Ad_g(\xi))$$  \hspace{1cm} (1.4)$$

for $\xi \in \mathfrak{g}$.

The identification of an element in $TG$ with an element in $G \times \mathfrak{g}$ via the left trivialization is also said to be the expression of $TG$ in **body coordinates** while the right trivialization is referred to as the expression of $TG$ in **space coordinates**. The terminology body and space is due to physical considerations of the rigid body, and has been expanded to the general setting of systems on Lie groups. We next briefly examine the setting of the classical three dimensional rigid body with $G = SO(3)$.

The motion of a rigid body can be viewed from a fixed point on the surface of the body, in which case the ambient space around the body will seem to be moving and other points on the body will seem fixed. Conversely, the motion can be viewed from the perspective of the absolute frame in space, in which case the body will rotate and change position over time while space is fixed. For example, consider a rotating carousel. To a child on the carousel (from the perspective of the moving frame), the other children on the carousel do not move, while the children waiting in line spin around. On the other hand, from the perspective a child waiting in line (absolute frame), the children on the carousel spin around while the other children waiting in line are stationary.

Let $R(t)$ denote the rotation of the moving frame as seen from the absolute frame at time $t$. In the case of the three dimensional free rigid body, $R(t) \in SO(3)$. A point on the body, $X$, has fixed coordinates with respect to the moving body frame, but varying coordinates with respect to the spatial frame. The so called space coordinates at time $t$ of the initial point $X$ on the body are given by

$$x(X, t) = R(t)X,$$  \hspace{1cm} (1.5)$$
from which we can compute the angular velocity with respect to space,

\[ \dot{x}(t) = \frac{\partial x(X,t)}{\partial t} = \dot{R}(t)X = \dot{R}(t)R^{-1}(t)x. \quad (1.6) \]

The angular velocity \( \dot{R}(t) \) is in general a vector in the tangent space \( T_{R(t)}(G) \), and the product \( \dot{R}(t)R^{-1}(t) \in g \) is a skew-symmetric matrix. We thus obtain the expression of the velocity with respect to coordinates in the fixed frame in space by multiplying on the right by \( R^{-1}(t) \), and the right trivialization is referred to as space coordinates for \( TG \).

Conversely, let \( x \) denote an initial point in space whose coordinates with respect to the moving body frame will change over time. When the body rotates in one direction, to a point on the body it appears as though ambient space is rotating in the opposite direction. We can thus express the position of the point \( x \) with respect to the body frame at time \( t \) by

\[ X(x,t) = R^{-1}(t)x. \quad (1.7) \]

In order to compute the angular velocity in body coordinates, we must account for the change in orientation of the motion (thus the minus sign), and we differentiate the expression for body coordinates

\[ \dot{X}(t) = -\frac{\partial X(x,t)}{\partial t} = R^{-1}(t)\dot{R}(t)R^{-1}(t)x = R^{-1}(t)\dot{R}(t)X. \quad (1.8) \]

Therefore the angular velocity with respect to coordinates in the body frame, is obtained by multiplying the angular velocity \( \dot{R}(t) \) on the left by \( R^{-1}(t) \).
Analogously, we can identify the cotangent bundle $T^*G$ with $G \times \mathfrak{g}^*$ in the following two ways:

$$\lambda^* : T^*G \to G \times \mathfrak{g}^* : \mu_g \mapsto (g, T_e^* L_g(\mu_g))$$  \hspace{1cm} (1.9)
$$\rho^* : T^*G \to G \times \mathfrak{g}^* : \mu_g \mapsto (g, T_e^* R_g(\mu_g))$$  \hspace{1cm} (1.10)

It is important to note a distinction between the tangent and cotangent trivializations. In the cotangent formulation we wish to identify a form $\mu_g \in T^*_g G$ with a form $\mu \in T^*_e G \cong \mathfrak{g}^*$. As always we must pull the form $\mu_g$ back into the cotangent space above the identity. Therefore, in the case of the left cotangent trivialization, we consider the map $L_g$ which maps the identity to the element $g$. The pull-back $T^*_e L_g$ pulls a form $\mu_g \in T^*_g G$ back to a form $\mu \in T^*_e G \cong \mathfrak{g}^*$. Again, we similarly call $\lambda^*(\mu_g)$ and $\rho^*(\mu_g)$ the expression of $\mu_g$ in body and space coordinates respectively. For $\mu \in \mathfrak{g}^*$, the transition from body to space coordinates is denoted

$$(\rho^* \circ (\lambda^*)^{-1}) : G \times \mathfrak{g}^* \to G \times \mathfrak{g}^* : (g, \mu) \mapsto (g, T_e^* R_g \circ T_g^* L_g^{-1}(\mu)) = (g, Ad_{g^{-1}}(\mu))$$  \hspace{1cm} (1.11)

1.1.2 Hamiltonian Vector Fields on the Cotangent Bundle

We next summarize how to translate the dynamics of a left invariant Hamiltonian vector field to the dual Lie algebra $\mathfrak{g}^*$. The same discussion holds for right invariant vector fields (using space coordinates for $T^*G$). For a complete treatment of this subject, see references [2], [1], or [24]. Recall there is a natural symplectic form on $T^*G$ and denote $H : T^*G \to \mathbb{R}$ a left-invariant Hamiltonian on $T^*G$ with corresponding left invariant Hamiltonian vector field $X_H$. We can use the left trivialization of $T^*G$ to express $X_H$ in body coordinates as follows,

$$T\lambda^*(X_H) : G \times \mathfrak{g}^* \to T(G \times \mathfrak{g}^*) : (g, \mu) \mapsto v_{(g,\mu)} = (g, \mu, v_g, v_\mu),$$  \hspace{1cm} (1.12)
where the vector \( v_{(g,\mu)} \in T(G \times g^*) \) can be decomposed into orthogonal components \( v_g \) and \( v_\mu \) which are tangent vectors to the group \( G \) and the dual Lie algebra \( g^* \) respectively.

Let \( \bar{X}_H : (g,\mu) \rightarrow T_\mu g^* \cong g^* \) denote the projection of the vector field \( T\lambda^*(X_H) \) (\( X_H \) in body coordinates) onto the component \( v_\mu \in g^* \). Due to the left invariance of the vector field \( X_H \), the vector field \( \bar{X}_H(g,\mu) \) is independent of the group \( G \), and therefore we can disregard from which point \( g \in G \) we started. The so called **cotangent Euler vector field**, \( \bar{Y} : g^* \rightarrow g^* \), is the expression of the left invariant vector field \( X_H \) in body coordinates. Thus it is sufficient to restrict the dynamics from the symplectic manifold \( T^*G \) to the dual Lie algebra \( g^* \), which has a Lie-Poisson structure, and the corresponding equations on \( g^* \) are described by the following proposition.

**Proposition 1.1.3** [24] Let \( \bar{H} = H|_{g^*} \) denote the restriction of a left invariant Hamiltonian to the dual Lie algebra \( g^* \). For any \( \mu \in g^* \), the restriction of \( X_H \) to \( g^* \) induced by the left trivialization of \( T^*G \) (cotangent Euler vector field), \( \bar{X} : g^* \rightarrow g^* \), can be expressed as follows:

\[
\bar{X}(\mu) = ad^{*}_{\bar{H}(\mu)}(\mu).
\] (1.13)

### 1.1.3 Coadjoint Orbit Reduction

Now we demonstrate how the dynamics in \( g^* \) given by Proposition 1.1.3 can be further reduced to the coadjoint orbits in \( g^* \) using a special case of symplectic reduction. Let \( \sigma_\mu = \{Ad^*_g\mu : g \in G\} \) denote the coadjoint orbit of \( \mu \in g^* \), and let \( \bar{\mu} \in \sigma_\mu \) be an arbitrary element in the coadjoint orbit of \( \mu \). Denote the pairing of \( g^* \) and \( g \) by \( (\mu,\xi) \) for any \( \mu \in g^* \) and \( \xi \in g \). We will first describe the symplectic structure of the coadjoint orbits by using the fundamental Kirillov-Kostant-Souriau form.
**Theorem 1.1.4** The tangent space of $\sigma_\mu$ at a point $\bar{\mu}$ can be characterized as

$$T_{\bar{\mu}}\sigma_\mu = \{ ad^*_{\xi} \bar{\mu} : \xi \in \mathfrak{g} \}.$$ \hspace{1cm} (1.14)

**Proof** Let $\xi \in \mathfrak{g}$ and let $g(t)$ be a curve in $G$ such that $g(0) = id$ and $g'(0) = \xi$. Then $\bar{\mu}(t) = Ad^*_{g(t)} \bar{\mu}$ is a curve in $\sigma_\mu$ with $\bar{\mu}(0) = \bar{\mu}$, and therefore for all $\eta \in \mathfrak{g}$ we have $(\bar{\mu}(t), \eta) = (\bar{\mu}, Ad_{g(t)}\eta)$. Taking the derivative of both sides we obtain

$$\frac{d}{dt} \bigg|_{t=0} (\bar{\mu}(t), \eta) = (\bar{\mu}'(0), \eta)$$

$$\frac{d}{dt} \bigg|_{t=0} (\bar{\mu}, Ad_{g(t)}\eta) = (\bar{\mu}, ad\xi\eta) = (ad^*_{\xi} \bar{\mu}, \eta).$$

Therefore any tangent vector $\bar{\mu}'(0) \in T_{\bar{\mu}}\sigma_\mu$ is of the form $ad^*_{\xi} \bar{\mu}$ for some $\xi \in \mathfrak{g}$. \hfill \Box

**Theorem 1.1.5** (Kirillov-Kostant-Souriau Theorem) Given any two tangent vectors $ad^*_{\xi} \bar{\mu}$ and $ad^*_{\eta} \bar{\mu}$ in $T_{\bar{\mu}}\sigma_\mu$, where $\xi, \eta \in \mathfrak{g}$, the coadjoint orbit $\sigma_\mu$ has a natural symplectic structure given by the formula

$$\omega_\mu \left( ad^*_{\xi} \bar{\mu}, ad^*_{\eta} \bar{\mu} \right) = - (\bar{\mu}, [\xi, \eta]).$$ \hspace{1cm} (1.15)

Now consider the restriction of the original left invariant Hamiltonian $H$ on $T^*G$ to the coadjoint orbits $\sigma_\mu$ in $\mathfrak{g}^*$, which we denote $\tilde{H} = H|_{\sigma_\mu}$. Using the Kirillov-Kostant-Souriau symplectic form defined on $\sigma_\mu$, we first find an expression for the Hamiltonian vector field of $\tilde{H}$, which we denote $\tilde{X}_{\tilde{H}}$. It will immediately follow that $\tilde{X}_{\tilde{H}}$ is the restriction of the cotangent Euler vector field $\tilde{X}_H$ (see (1.13)) to the coadjoint orbits, and thus the solutions of $\tilde{X}_H$ and $\tilde{X}_{\tilde{H}}$ are identical.

**Theorem 1.1.6** Let $\tilde{H}$ denote the restriction of a left invariant Hamiltonian $H$ on $T^*G$ to a coadjoint orbit $\sigma_\mu$ in $\mathfrak{g}^*$. The Hamiltonian vector field of $\tilde{H}$ can be
expressed as

\[ \tilde{X}_H(\bar{\mu}) = ad^{\ast}_{d\tilde{H}(\bar{\mu})}\bar{\mu}. \] (1.16)

**Proof** For a given \( \bar{\mu} \in \sigma_\mu \), clearly \( \tilde{X}_H(\bar{\mu}) \in T_\mu \sigma_\mu \), and thus \( \tilde{X}_H(\bar{\mu}) = ad^{\ast}_\xi \bar{\mu} \) for some \( \xi \in \mathfrak{g} \) (by Theorem 1.1.4). Using the KKS structure on \( \sigma_\mu \), we know for any \( ad^{\ast}_\eta \bar{\mu} \in T_\mu \sigma_\mu \) that

\[
\omega_\mu \left( \tilde{X}_H(\bar{\mu}), ad^{\ast}_\eta \bar{\mu} \right) = \left( ad^{\ast}_\eta \bar{\mu}, d\tilde{H}(\bar{\mu}) \right) = \left( \bar{\mu}, ad_\eta \left( d\tilde{H}(\bar{\mu}) \right) \right) = \left( \bar{\mu}, [\eta, d\tilde{H}(\bar{\mu})] \right)
\]

and using the symplectic form defined by the KKS theorem in (1.1.5)

\[
\omega_\mu \left( \tilde{X}_H(\bar{\mu}), ad^{\ast}_\eta \bar{\mu} \right) = \omega_\mu \left( ad^{\ast}_\xi \bar{\mu}, ad^{\ast}_\eta \bar{\mu} \right) = - (\bar{\mu}, [\xi, \eta]) = (\bar{\mu}, [\eta, \xi]),
\]

from which it can be concluded that \( \xi = d\tilde{H}(\bar{\mu}) \in \mathfrak{g} \) and therefore we express the restriction of the Hamiltonian vector field as \( \tilde{X}_H(\bar{\mu}) = ad^{\ast}_{d\tilde{H}(\bar{\mu})}\bar{\mu} \).

**Corollary 1.1.7** The cotangent Euler vector field \( \tilde{X}_H \) in (1.13) is Hamiltonian on the coadjoint orbits with Hamiltonian \( \tilde{H} = H|_{\sigma_\mu} \).

Reviewing the reduction process for Lie group systems with symmetry, the reduction occurs in two stages. First we use the symmetry to translate the dynamics over the identity into the dual Lie algebra \( \mathfrak{g}^\ast \). The reduction onto the Poisson system in the dual Lie algebra is described by the cotangent Euler vector field \( \tilde{X}_H \) in (1.13). The cotangent Euler vector field then projects onto Hamiltonian system \( \tilde{X}_H \) on the coadjoint orbits with Hamiltonian \( \tilde{H} \).

1.1.4 General Symplectic Reduction

We note that the coadjoint orbit reduction process for systems on Lie groups is a result of Arnold’s work with the classical Euler equation for the flow of an ideal
fluid. However, many similar systems with symmetries which allow the dynamics to be reduced do not have configuration spaces which are Lie groups. Marsden and Weinstein discover the principle of symplectic reduction for these general systems whose phase space is a symplectic manifold $P = T^*Q$, where $Q$ is now a general manifold [22]. We summarize their general process below:

Let $(P, \omega, \Phi, J)$ denote a symplectic manifold $P$ with a $G$-action $\Phi : G \times P \to P$ which has an equivariant moment map $J$. For a regular value $\mu \in g^*$ of $J$, consider the isotropy subgroup $G_\mu$ of $G$. Since $J$ is equivariant, $G_\mu$ is a well-defined, free, and proper action on $J^{-1}(\mu)$, and thus the quotient $P_\mu = \frac{J^{-1}(\mu)}{G_\mu}$ is a smooth submanifold of $P$.

**Proposition 1.1.8** [22] Denote the natural projection $\pi_\mu : J^{-1}(\mu) \to P_\mu$. The smooth submanifold $P_\mu$ has a unique symplectic structure $\tilde{\omega}_\mu$ which is derived from the original symplectic structure $\omega$ on $P$ in the following way

$$\pi_\mu^* \tilde{\omega}_\mu = i_\mu^* \omega$$

(1.17)

where $i_\mu^*$ is the pull-back of the canonical inclusion $i_\mu : J^{-1}(\mu) \to P$. The manifold $P_\mu$ is called a reduced symplectic manifold.

Given a Hamiltonian system $X_H$ on the symplectic manifold $P$, it is true in general that dynamics of the Hamiltonian system $X_H$ on $P$ can be reduced to simpler Hamiltonian systems on the reduced submanifolds $P_\mu$. The solutions in $P_\mu$ can then be reconstructed using the symmetry of the system to describe the complete dynamics on $P$.

1.1.5 The Inertia Map

Generally speaking, the elements of $g^*$ are linear functionals on the Lie algebra $g$, and thus finding explicit equations of motion for the reduced Hamiltonian system
in the coadjoint orbits presents an added challenge we aim to avoid. The elements
of the Lie algebra are more tangible objects, and in the case of the rigid body
they are skew-symmetric matrices. For this reason, we now wish to develop similar
symplectic reduction in the tangent bundle as performed in the previous section
with the cotangent bundle. In general if we can introduce a bilinear, symmetric,
non-degenerate, Ad-invariant form on \( g \), then it is possible to pass the symplectic
structure of the coadjoint orbits in \( g^* \) to the adjoint orbits in \( g \). For a semi-simple
Lie algebra such as \( \mathfrak{so}(n) \), the Killing form will provide such a form which we will
then use to find explicit equations for the reduced Hamiltonian system in adjoint
orbits of the Lie algebra \( g \).

Let \( < , > \) denote a left-invariant metric on \( G \), and let \( E : TG \to \mathbb{R} \) be an
arbitrary left-invariant smooth energy (quadratic) function. The metric \( < , > \)
induces an isomorphism between \( TG \) and \( T^*G \) via the identification
\[
\tau : TG \to T^*G : v_g \mapsto \mu_g
\]
where \( \mu_g(w_g) = < v_g, w_g > \) for all \( g \in G \) and \( w_g \in T_gG \). This identification yields
a symplectic structure \( \omega' = \tau^*(\omega) \) on \( TG \) which is the pull-back of the symplectic
structure \( \omega \) on \( T^*G \). Now given the corresponding left invariant Hamiltonian vector
field \( X_E \) on \( TG \), we can similarly express this vector field in body coordinates as pre-
viously described in the cotangent case. The expression of \( X_E \) in body coordinates
on \( g \) is analogously called the tangent Euler vector field.

**Remark 1.1.9** Provided \( E \) is a smooth, left invariant, quadratic function, the flows
of the tangent Euler vector field of \( X_E \) are the geodesic sprays of the metric
\( < , > \). This fact follows from the equivalence of the principle of least action and
the length minimizing property of geodesics. Therefore the cotangent Euler field of
\( X_E \) is uniquely determined by the choice of left invariant metric \( < , > \) on \( TG \).
Since the tangent Euler vector field of \( X_E \) is the same for all smooth, left invariant,
energy functions $E$ on $TG$, we can define the energy function by left translating $K : \mathfrak{g} \to \mathbb{R} : \xi \mapsto \frac{1}{2} \langle \xi, \xi \rangle$. Thus the metric $< \cdot, \cdot >$ is referred to as the energy metric. However, if $E$ is not quadratic then it does not necessarily follow that the flows of $X_E$ are geodesic sprays of some metric $< \cdot, \cdot >$.

Next consider the special case when $\mathfrak{g}$ has a second bilinear, symmetric, non-degenerate form $(\cdot, \cdot)$ which is invariant under the $Ad$-action and thus satisfies

$$<(Ad_g \eta, Ad_g \xi) = (\xi, \eta)> \tag{1.19}$$

for any $\xi$ and $\eta \in \mathfrak{g}$ and all $g \in G$. In such a situation we can map the symplectic structure of the coadjoint orbits in $\mathfrak{g}^*$ to the adjoint orbits in $\mathfrak{g}$.

As with the energy metric $(\cdot, \cdot)$ in (1.18), this second non-degenerate form $(\cdot, \cdot)$ induces another isomorphism between the $\mathfrak{g}^*$ and $\mathfrak{g}$. We identify each $\xi \in \mathfrak{g}$ with a linear functional

$$f_\xi : \mathfrak{g} \to \mathbb{R} : \eta \mapsto (\xi, \eta). \tag{1.20}$$

For any $\mu$ and $\eta$ in $\mathfrak{g}^*$ and $\mathfrak{g}$ respectively, we denote the pairing of $\mathfrak{g}^*$ with $\mathfrak{g}$ induced by the $Ad$-invariant form $(\cdot, \cdot)$ by

$$((\mu, \eta)) = (\xi, \eta) \quad \text{for the unique} \quad \xi \in \mathfrak{g} \quad \text{such that} \quad \mu = f_\xi. \tag{1.21}$$

**Theorem 1.1.10** [24] Let $G$ be a Lie group endowed with a left invariant metric denoted $< \cdot, \cdot >$. If $G$ admits a second bilinear, symmetric, non-degenerate form $(\cdot, \cdot)$ which is invariant under the Adjoint action of $G$, then there exists a unique, linear, $(\cdot, \cdot)$ symmetric (meaning $((\tilde{J}(\xi), \eta)) = ((\tilde{J}(\eta), \xi))$ for all $\xi, \eta \in \mathfrak{g}$), positive isomorphism $\tilde{J} : \mathfrak{g} \to \mathfrak{g}^*$ such that the two forms above are related by the property

$$((\tilde{J} \cdot, \cdot)) =< \cdot, \cdot >. \tag{1.22}$$
Figure 1.1: Body and space coordinates for a general Lie group $G$. The inertia map $\tilde{J}$ connects the Ad-invariant induced pairing of $\mathfrak{g}$ and $\mathfrak{g}^*$ denoted $(\cdot, \cdot)$ (for example the Killing form) with the left invariant energy metric denoted $\langle \cdot, \cdot \rangle$. Since the energy metric is left invariant, the inertia map $\tilde{J}$ commutes with the left trivializations of $TG$ and $T^*G$, but not necessarily with the right trivializations. The map $\tilde{J}_g$ denotes the left translation of the inertia map $\tilde{J}$ from the identity to the tangent and cotangent spaces over an arbitrary $g \in G$.

Essentially the map $\tilde{J}$ is the Legendre transform which moves the equations of motion from the Lagrangian formulation with respect to the energy metric $\langle \cdot, \cdot \rangle$ in $TG$ to the Hamiltonian formulation with respect to the natural (Killing form induced) pairing in $T^*G$. For this reason the map $\tilde{J}$ is referred to as the inertia map. In the case of the rigid body, $\tilde{J}$ maps angular velocities in $\mathfrak{g}$ to angular momenta in $\mathfrak{g}^*$. Physically speaking, $\tilde{J}$ describes the mass distribution of the rigid body.

1.1.6 Identifying Coadjoint and Adjoint Orbits

In the definition of the inertia map $\tilde{J}$, we have been careful to distinguish linear functions in $\mathfrak{g}^*$ from matrices in $\mathfrak{g}$ since generally speaking the two are not isomorphic. However, we now wish to use the Ad-invariant identification of $\mathfrak{g}$ and $\mathfrak{g}^*$ to
move the reduced equations of $\tilde{X}_H$ obtained in (1.13) from the coadjoint orbits in $\mathfrak{g}^*$ to adjoint orbits in $\mathfrak{g}$. First notice that after the identification of $\mathfrak{g}^*$ and $\mathfrak{g}$ induced by the form $(\ , \ )$ on $\mathfrak{g}$, the inertia map in (1.22) can be defined as the unique, linear, $(\ , \ )$-symmetric, positive isomorphism from $\mathfrak{g}$ into itself since

$$(\tilde{J}(\xi), \eta) = ((\tilde{J}(f_\xi), \eta)) = \langle \xi, \eta \rangle$$

(1.24)

for all $\xi$ and $\eta$ in $\mathfrak{g}$.

The equivariant isomorphism induced by $(\ , \ )$ moreover maps the KKS symplectic structure (1.15) of coadjoint orbits to an analogous symplectic structure on adjoint orbits:

$$(\omega_\xi (\text{ad}_\eta \tilde{\xi}, \text{ad}_\xi \tilde{\eta}) = - (\{\eta, \xi\}, \tilde{\xi})$$

(1.25)

where the tangent space at $\tilde{\xi} \in \sigma_\xi$ is characterized by $T_{\xi}\sigma_\xi = \{\text{ad}_\eta \tilde{\xi} : \eta \in \mathfrak{g}\}$ [24]. Finally using the induced symplectic structure (1.25) on the adjoint orbits, we obtain an analogous reduced system (as in Theorem 1.1.6) now defined more concretely on the adjoint orbits.

**Theorem 1.1.11** [24] Let $G$ be a Lie group with left invariant energy metric $<\ , \ >$, Ad-invariant (Killing) form $(\ , \ )$, and inertia map $\tilde{J}$. Consider the the smooth energy function $\tilde{K}(\xi) = \frac{1}{2} <\xi, \xi> = \frac{1}{2} (\tilde{J}(\xi), \xi)$ (defined on $\mathfrak{g}$ and left translated to a function $K$ on $TG$), and denote $\tilde{K} = K|_{\sigma_\xi}$ the restriction of $K$ to an adjoint orbit $\sigma_\xi$. The vector field $\tilde{X}_K$ (the left trivialization of the vector field $X_K$ on $TG$) is Hamiltonian on the adjoint orbits of $\sigma_\xi$, with Hamiltonian $\tilde{K}$. In particular, the flow of $\tilde{X}_K$ through a point $\tilde{\xi} \in \sigma_\xi$ satisfies the following equation:

$$X_{\tilde{K}} \left(\tilde{J}(\tilde{\xi})\right) = [\tilde{J}(\tilde{\xi}), \tilde{\xi}].$$

(1.26)

In order to study the dynamics of $X_K$ in $TG$, we can similarly use the symmetry of the system to find the reduced equation $\tilde{X}_K$ in $\mathfrak{g}$. Provided $\mathfrak{g}$ has a non-degenerate,
Ad-invariant form, the Lie algebra $\mathfrak{g}$ has an induced Lie-Poisson structure which can then be used to further reduce the system onto the adjoint orbits.

1.2 Deriving Euler’s Equation

Theorem 1.1.11 is a beautiful consequence of the abundant geometry underlying systems on Lie groups. In this section we derive the equations of motion of the $n$ dimensional rigid body by applying Theorem 1.1.11 to the special case when the Lie group $G = SO(n)$. We will first identify the dual Lie algebra $\mathfrak{g}^* = \mathfrak{so}^*(n)$ with the Lie algebra $\mathfrak{g} = \mathfrak{so}(n)$. Using body coordinates for the rigid body, we then construct the equations of motion in terms of body angular velocities $\Omega \in \mathfrak{g}$ (1.8). In order to construct the equations of motion, we need to define the energy metric of the system and the corresponding inertia map $\tilde{J}$. Euler’s equation will be an immediate consequence of Theorem 1.1.11. Lastly we look at the familiar three dimensional rigid body and verify the generalized Euler’s equation for the $n$ dimensional rigid body is indeed consistent with Euler’s classical expression for the three dimensional rigid body:

\[
\begin{align*}
I_1 \dot{\omega}_1 &= (\lambda_2 - \lambda_3) \omega_2 \omega_3 \\
I_2 \dot{\omega}_2 &= (\lambda_3 - \lambda_1) \omega_1 \omega_3 \\
I_3 \dot{\omega}_3 &= (\lambda_1 - \lambda_2) \omega_1 \omega_2 
\end{align*}
\] (1.27)

1.2.1 Identifying Angular Velocity and Angular Momentum

The Lie algebra $\mathfrak{g} = \mathfrak{so}(n)$ is a semi-simple Lie algebra, and thus the Killing form, $B(X, Y) = \text{tr}(ad_X, ad_Y)$, is negative definite. The Killing form is a non-degenerate, Ad-invariant form which induces a natural identification of skew symmetric matrices in $\mathfrak{g}$ with linear functionals in the dual Lie algebra $\mathfrak{g}^*$. In particular, consider the following form,

\[
(X, Y) := -\frac{1}{2} \text{tr}(XY) \quad \text{for all } X \text{ and } Y \in \mathfrak{so}(n),
\] (1.28)
a scalar multiple (which depends on $n$) of the Killing form.

Given the form $(\cdot, \cdot)$ and a linear functional $\mu \in \mathfrak{so}^*(n)$, we can now represent $\mu$ as a skew symmetric matrix in $\mathfrak{so}(n)$ by choosing the unique $X \in \mathfrak{so}(n)$ such that $\mu = f_X = (X, \cdot)$. Strictly speaking, skew-symmetric matrices in $\mathfrak{so}(n)$ are angular velocities, linear functionals in $\mathfrak{so}(n)$ are angular momentum, and the inertia map $\bar{J} : \mathfrak{g} \to \mathfrak{g}^*$ converts velocities into momenta; however, having washed our hands of this distinction between elements of $\mathfrak{so}(n)$ and $\mathfrak{so}^*(n)$, from now on we work only in the Lie algebra $\mathfrak{g} = \mathfrak{so}(n)$. Angular momenta are represented as skew-symmetric matrices which live in the Lie algebra $\mathfrak{so}(n)$, and we can conveniently express the equations of motion for the rigid body more concretely in terms of matrices.

1.2.2 The Energy Metric

Let $R(t) \in SO(n)$ denote a rotation in $\mathbb{R}^n$, and thus $R(t)x$ describes the position of a particle at time $t$ initially located at $x \in \mathbb{R}^n$. Let $\mu$ be a positive measure on $\mathbb{R}^n$ which describes the mass distribution of the body. In particular, $\mu$ is assumed not to have support which is a one dimensional subspace. The kinetic energy for a rigid body rotating freely about the center of mass is given by the integral

$$H(t) = \frac{1}{2} \int_{\mathbb{R}^n} \|\tilde{\Omega}(t)x\|^2 \ d\mu(x)$$

where $\|\cdot\|$ denotes the standard Euclidean norm in $\mathbb{R}^n$. Let $\Omega_S(t) = \rho(\tilde{R}(t)) = \tilde{R}(t)R^{-1}(t) \in \mathfrak{g}$ denote the angular velocity in space coordinates which is translated to $\mathfrak{g}$ from the right (1.6). Since $R(t) \in SO(n)$, it preserves the inner product, and
we can simplify the below

\[ \| \hat{R}(t) x \|^2 = \| \hat{R}(t) R^{-1}(t) R(t) x \|^2 \]
\[ = \| \Omega_S(t) R(t) x \|^2 \]
\[ = \| R^{-1}(t) \Omega_S(t) R(t) x \|^2 \]
\[ = \| (Ad_{R^{-1}(t)}(\Omega_S(t))) x \|^2 \]
\[ = \| \Omega_B(t) x \|^2 \]  \hspace{1cm} (1.30)

where \( \Omega_B(t) = Ad_{R^{-1}(t)}(\Omega_S(t)) = R^{-1}(t) \hat{R}(t) \in \mathfrak{g} \) is body angular velocity translated to \( \mathfrak{g} \) from the left (1.8).

Since the norm above is left invariant, the energy function \( H \) is similarly left invariant. Therefore, we wish to restrict the function \( H \) to the Lie algebra \( \mathfrak{g} \) by left trivialization. Since we now exclusively use body coordinates for \( TG \), we omit the subscript \( B \) in the expression for the body angular velocity \( \Omega_B(t) \). The angular velocity with respect to the moving body frame is thus simply denoted \( \Omega(t) \). The energy \( H \) is thus expressed in terms of body coordinates on the Lie algebra \( \mathfrak{g} \) as follows,

\[ \tilde{H}(\Omega(t)) = \frac{1}{2} \int_{\mathbb{R}^n} \| \Omega(t) x \|^2 \, d\mu(x) = \frac{1}{2} \langle \Omega(t), \Omega(t) \rangle, \]  \hspace{1cm} (1.31)

where \( <,> \) is the left invariant energy metric on \( TG \). Since for the rest of the paper we concern ourselves only with understanding the dynamics of the reduced Lie-Poisson system in \( \mathfrak{so}(n) \), we no longer use the \( \sim \) to distinguish the energy function on \( \mathfrak{g} \) from the unreduced energy function \( H \) on \( TG \). From now the function \( H \) refers to the restriction of the energy to \( \mathfrak{g} \), and we no longer bother using the \( \sim \) notation.
1.2.3 The Inertia Map of the Rigid Body

We denote the inertia map $\tilde{J}_0 : \mathfrak{so}(n) \to \mathfrak{so}(n)$, and recall it is the unique linear map which satisfies the identity

$$(\tilde{J}_0(A), B) = -\frac{1}{2} \text{tr} \left( (\tilde{J}_0(A)B) = \langle A, B \rangle = \int_{\mathbb{R}^n} A \cdot B \, d\mu(x). \right) \quad (1.32)$$

For $1 \leq i, j \leq n$, let $e_{ij}$ denote the standard basis element for $\text{Mat}_{n \times n}(\mathbb{R})$, and therefore the set $\{M_{ij} = e_{ij} - e_{ji} | 1 \leq i < j \leq n\}$ is a basis for $\mathfrak{so}(n)$. Let $\tilde{J}_0(A)_{ij}$ denote the $ij^{th}$ entry of $\tilde{J}_0(A)$. Since $\tilde{J}_0(A) \in \mathfrak{g}$, it follows $\tilde{J}_0(A)_{ij} = -\tilde{J}_0(A)_{ji}$.

Setting $B = M_{ij}$, we simplify the left hand side of (1.32)

$$(\tilde{J}_0(A), M_{ij}) = \left( \tilde{J}_0(A), e_{ij} \right) - \left( \tilde{J}_0(A), e_{ji} \right)$$

$$= -\frac{1}{2} \text{tr} \left( \tilde{J}_0(A) \cdot e_{ij} \right) + \frac{1}{2} \text{tr} \left( \tilde{J}_0(A) \cdot e_{ji} \right)$$

$$= \frac{1}{2} \left( \tilde{J}_0(A)_{ji} \right) - \frac{1}{2} \left( \tilde{J}_0(A)_{ij} \right)$$

$$= \tilde{J}_0(A)_{ji}. \quad (1.33)$$

Let $x = (x_1, \ldots, x_n)$ and define $c_{ij} = \int_{\mathbb{R}^n} x_i x_j \, d\mu(x)$. We next simplify the right hand side of (1.32) when $B = M_{ij}$:

$$\langle A, M_{ij} \rangle = \left( \int_{\mathbb{R}^n} A \cdot e_{ij} x \, d\mu(x) \right) - \left( \int_{\mathbb{R}^n} A \cdot e_{ji} x \, d\mu(x) \right)$$

$$= \int_{\mathbb{R}^n} \sum_{k=1}^{n} (a_{ik} x_k x_j) \, d\mu(x) - \int_{\mathbb{R}^n} \sum_{k=1}^{n} (a_{jk} x_k x_i) \, d\mu(x)$$

$$= \sum_{k=1}^{n} a_{ik} c_{kj} - \sum_{k=1}^{n} a_{jk} c_{ki}$$

$$= (CA + AC)_{ij} \quad (1.34)$$

where $C$ is the $n \times n$ symmetric matrix with $ij^{th}$ entry equal $c_{ij} = \int_{\mathbb{R}^n} x_i x_j \, d\mu(x) =$
Thus by equations (1.33) and (1.34) we have proved

\[ \tilde{J}_0A = CA + AC \] [24].

(1.35)

Since \( C \) is a real, symmetric matrix, it can be diagonalized. Therefore, there exists a matrix \( g \) in \( SO(n) \) such that \( gCg^{-1} = D \), where \( D \) is a diagonal matrix. We wish to introduce a new basis for \( \mathbb{R}^n \) such that the inertia map is diagonalized in the sense \( \tilde{J}(A) = DA + AD \) for \( D \) diagonal.

Consider a new positive measure on \( \mathbb{R}^n \) obtained from the original measure \( \mu \) by \( \nu(x) = \mu(g^{-1}x) \), where \( g \in SO(n) \) is the matrix which diagonalizes \( C \) in (1.34). Now consider a modification of the inertia map \( \tilde{J}_0 \) in (1.35) defined by \( \tilde{J}(A) = g \left( \tilde{J}_0(g^{-1}Ag) \right) g^{-1} \), which we prove in the next theorem to be the corresponding inertia map with respect to new energy form \( < , > \) expressed with respect to the measure \( \nu \) on \( \mathbb{R}^n \).

**Theorem 1.2.1** The map \( \tilde{J}(A) = DA + AD \), and it satisfies the definition of the inertia map (1.24), namely for all \( A, B \in \mathfrak{g} \)

\[ -\frac{1}{2} \left( \tilde{J}(A), B \right) = \int_{\mathbb{R}^n} A y \cdot B y \ d\nu(y). \] [24]

**Proof** We simplify the following expression to prove the first claim:

\[
\tilde{J}(A) = g \left( \tilde{J}_0(g^{-1}Ag) \right) g^{-1} \\
= g \left( Cg^{-1}Ag \right) g^{-1} + g \left( g^{-1}AgC \right) g^{-1} \\
= \left( gCg^{-1} \right) A + A \left( gCg^{-1} \right) \\
= DA + AD.
\]

Let \( y = gx \). Using the Ad-invariance of \( < , > \) along with the left invariance of \( < , > \),
we simplify the necessary condition for the inertia map:

\[
\left( \tilde{J}(A), B \right) = \left( g \left( \tilde{J}_0(g^{-1}Ag) \right) g^{-1}, B \right) = \left( \tilde{J}_0(g^{-1}Ag), g^{-1}Bg \right) = \left\langle g^{-1}Ag, g^{-1}Bg \right\rangle = \left\langle Ag, Bg \right\rangle = \int_{\mathbb{R}^n} A(gx) \cdot B(gx) \, d\mu(x) = \int_{\mathbb{R}^n} Ay \cdot By \, d\nu(y).
\]

Depending on the mass distribution of the body, we thus choose a new orthonormal basis for \( \mathbb{R}^n \) (namely \( y = gx \)) for which the corresponding inertia map \( \tilde{J} \) is of the form \( \tilde{J}(a) = DA + AD \) for \( D \) a real, diagonal \( n \times n \) matrix. From now on, we denote the diagonalized inertia map as follows:

\[
\tilde{J}(A) = JA + AJ \quad \text{for} \quad J = diag(\lambda_1, \lambda_2, \ldots, \lambda_n).
\]  

The diagonalized operator \( \tilde{J} \) above has eigenvalues \( \lambda_i + \lambda_j \) \((i < j)\) with the corresponding basis of eigenvectors \( M_{ij} = e_{ij} - e_{ji} \). At this point there are no further restrictions on the \( \lambda_i \)'s of \( J \) other than \( \lambda_i > 0 \) for all \( 1 \leq i \leq n \) since the map \( \tilde{J} \) is positive and the eigenvalues must therefore all be positive. The eigenvalues \( \lambda_i + \lambda_j \) are referred to as the principal moments of inertia, and they describe the mass distribution of the rigid body. We call the diagonal matrix \( J \) the inertia matrix. Given an angular velocity (with respect to the body) \( \Omega \), the matrix \( \tilde{J}(\Omega) = J\Omega + \Omega J = M \) is the corresponding angular momentum (in body coordi-
1.2.4 The Equations of Motion

**Theorem 1.2.2** Given a rigid body in \( \mathbb{R}^n \) there exists an orthonormal basis of \( \mathbb{R}^n \) which is determined by the mass distribution of the rigid body in which the equations of motion of the free rigid body about a fixed point have the form

\[
\dot{M} = [M, \Omega] \quad (1.37)
\]

for \( \Omega \) the angular velocity in body coordinates, \( M = J\Omega + \Omega J \) the corresponding angular momentum, and \( J \) the diagonal inertia matrix. These equations are Hamiltonian with respect to the function

\[
H : g \to \mathbb{R} : \Omega \mapsto -\frac{1}{2} \text{tr} (J\Omega^2) \quad . \quad (1.38)
\]

**Proof** This theorem is a direct result of Theorem 1.1.11 at the end of Section 1.1 with the left invariant energy metric defined in (1.31), the bilinear, symmetric, non-degenerate, Ad-invariant form defined in (1.28), and the diagonalized inertia map \( \bar{J} \) obtained in (1.36). Strictly speaking, Theorem 1.1.11 proves the corresponding Hamiltonian is \( H(\Omega) = \frac{1}{2} \langle \Omega, \Omega \rangle \); however, we can quickly verify this is equivalent to the Hamiltonian in (1.38):

\[
H(\Omega) = \frac{1}{2} \langle \Omega, \Omega \rangle \\
= \frac{1}{2} \left( \bar{J}(\Omega), \Omega \right) \\
= -\frac{1}{4} \text{tr} (J\Omega^2 + \Omega J\Omega) \\
= -\frac{1}{2} \text{tr} (J\Omega^2) \quad .
\]
1.2.5 Example: The 3 Dimensional Case

The equation of motion $\ddot{M} = [M, \Omega]$ is called the generalized Euler equation and was first constructed by Arnold in 1966 [2]. Euler originally discovered the equations of motion for the three dimensional rigid body around 1750:

\[
\begin{align*}
I_1 \dot{\omega}_1 &= (\lambda_2 - \lambda_3)\omega_2 \omega_3 \\
I_2 \dot{\omega}_2 &= (\lambda_3 - \lambda_1)\omega_1 \omega_3 \\
I_3 \dot{\omega}_3 &= (\lambda_1 - \lambda_2)\omega_1 \omega_2.
\end{align*}
\] (1.39)

We now study the special case of the three dimensional rigid body. This particular case is special for physical reasons, namely we can actually picture a body rotating in our three dimensional world. This case is also special for mathematical reasons as the corresponding Lie algebra $\mathfrak{so}(3)$ is isomorphic (as Lie algebras) to $\mathbb{R}^3$ with Lie bracket the vector cross product. Moreover this is the only case in which the dimension of the Lie group and Lie algebra are equal to the dimension of the space in which the body is rotating. Using the usual isomorphism of $\mathfrak{so}(3)$ with $\mathbb{R}^3$, we now prove that Euler’s classical equation (1.39) is consistent with the equations of motion for the generalized rigid body in (1.37).

First consider the basis for the Lie algebra $\mathfrak{g} = \mathfrak{so}(3)$: $X_1 = e_{32} - e_{23}$, $X_2 = e_{13} - e_{31}$, and $X_3 = e_{21} - e_{12}$. Using the canonical basis elements $e_1$, $e_2$, and $e_3$ for $\mathbb{R}^3$, we define a Lie bracket isomorphism between $\mathfrak{so}(3)$ and $\mathbb{R}^3$ as follows

\[
\mathfrak{so}(3) \to \mathbb{R}^3 : \begin{cases}
X_i \mapsto e_i & \text{for } i = 1, 2, 3 \\
[A, B] \mapsto a \times b
\end{cases}.
\] (1.40)

Using the isomorphism in (1.40) it follows $(A, B) = -\frac{1}{2} \text{tr}(AB) = a \cdot b$ (thus the convenience of the $-\frac{1}{2}$), and the inertia map $\bar{J}$ with corresponding matrix
$J = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$ is mapped to the linear map represented by the matrix $I$ below,

$$
\tilde{J}(\Omega) = J\Omega + \Omega J \rightsquigarrow I(\omega) = \begin{pmatrix}
\lambda_2 + \lambda_3 & 0 & 0 \\
0 & \lambda_1 + \lambda_3 & 0 \\
0 & 0 & \lambda_1 + \lambda_3
\end{pmatrix} \begin{pmatrix}
\omega_1 \\
\omega_2 \\
\omega_3
\end{pmatrix},
$$

and notice the eigenvalues of $I$ are similarly $I_1 = \lambda_2 + \lambda_3$, $I_2 = \lambda_1 + \lambda_3$, and $I_3 = \lambda_1 + \lambda_2$ with corresponding eigenvectors $e_1$, $e_2$, and $e_3$, respectively. By introducing a change of basis determined by the mass distribution as in Theorem 1.2.1, we have aligned the eigenvalues of the principal moments of inertia with the canonical axes in $\mathbb{R}^3$.

Next consider the angular velocity $\Omega \in \mathfrak{so}(3)$ which has corresponding vector $\omega \in \mathbb{R}^3$, along with the angular momentum $J\Omega + \Omega J = M$ identified with $I \omega \in \mathbb{R}^3$. Applying Theorem 1.2.2 in $\mathbb{R}^3$ we obtain the equivalent equation of motion

$$(\tilde{J}\Omega) \dot{} = [\tilde{J}\Omega, \Omega] \rightsquigarrow (I \omega) \times \omega,$$

and indeed we recover Euler’s classical equation

$$
\begin{align*}
I_1\dot{\omega}_1 &= (\lambda_2 - \lambda_3)\omega_2\omega_3 \\
I_2\dot{\omega}_2 &= (\lambda_3 - \lambda_1)\omega_1\omega_3 \\
I_3\dot{\omega}_3 &= (\lambda_1 - \lambda_2)\omega_1\omega_2.
\end{align*}
$$

The corresponding Hamiltonian is expressed in $\mathbb{R}^3$ by

$$
H(\omega) = \frac{1}{2} ((I \omega) \cdot \omega) = \frac{1}{2} \sum_{j=1}^{3} \frac{\omega_j^2}{I_j}.
$$

(1.41)

The level sets of $H$ are energy ellipsoids in $\mathbb{R}^3$, and the adjoint orbits are spheres centered at the origin. The solutions of the three dimensional rigid body are thus intersections of energy ellipsoids and adjoint orbits. Notice in the figure that the
equilibria are rotations about the axes which point on the same direction as the principal axes of inertia.

Figure 1.2: The trajectories of the three dimensional rigid body on an adjoint orbit.

1.3 Integrability

In this section, we prove the Euler’s equation for the $n$ dimensional rigid body are integrable. Moreover, we show that equation (1.37) is a restriction to $\mathfrak{so}(n)$ of the Dubrovin equation defined on $\mathfrak{g} = \mathfrak{sl}(n)$:

$$\begin{align*}
[A, V] = [[A, V], [B, V]].
\end{align*}$$

(1.42)

where $A = \text{diag}(a_1, a_2, \ldots, a_n)$, $B = \text{diag}(b_1, b_2, \ldots, b_n)$, and $V$ is an $n \times n$ matrix with zeros along the diagonal. Notice that for $V$ restricted to $\mathfrak{so}(n)$, since $[A, V]$ and $[B, V]$ are both skew symmetric, the Dubrovin equation leaves $\mathfrak{so}(n)$ invariant. Manakov sets $A = J^2$ and $B = J$ in order to recover Euler’s equation, and he generates a complete set of integrals which he proves are in involution, thus proving the $n$ dimensional rigid body is integrable [18].

**Definition 1.3.1** A Hamiltonian system $H$ of dimension $2N$ is an integrable system if there exist functions $F_k$ for $k = 1, 2, \ldots, N$ such that
1. Each $F_k$ is constant along the flows of $H$, or in other words $\{F_k, H\} = 0$.

2. The functions $F_k$ are in involution with each other. Thus fixing any $F_j$, all other functions $F_i$ are constant along the flows of $F_j$, or $\{F_j, F_i\} = 0$.

3. The corresponding differentials $dF_k$ form a linear independent set.

Since the flows of the Hamiltonian system corresponding to the rigid body are contained in the adjoint orbits of $\mathfrak{so}(n)$, it is sufficient to prove integrability on the adjoint orbits. The dimension of a non-degenerate adjoint orbit of a semi-simple Lie algebra $\mathfrak{g}$, such as $\mathfrak{so}(n)$, is equal to $\dim(\mathfrak{g}) - \text{rank}(\mathfrak{g})$. Therefore, the dimension of each generic adjoint orbit of $\mathfrak{so}(n) = \dim(\mathfrak{so}(n)) - \text{rank}(\mathfrak{so}(n))$, where $\dim(\mathfrak{so}(n)) = \frac{n(n-1)}{2}$ and $\text{rank}(\mathfrak{so}(n)) = \frac{n}{2}$ if $n$ is even or $\frac{n-1}{2}$ if $n$ is odd. It follows that the dimension of each non-degenerate adjoint orbit $\mathfrak{so}(n)$ is $\frac{n(n-2)}{2}$ if $n$ is even and $\frac{(n-1)^2}{2}$ if $n$ is odd. Therefore, we need to find $\frac{n^2-2n}{4}$ integrals when $n$ is even and $\frac{(n-1)^2}{4}$ integrals when $n$ is odd.

Since the three dimensional free rigid body is a two dimensional Hamiltonian system, it immediately follows that the system is integrable since the Hamiltonian is a conserved quantity. In order to show that a generic $n$ dimensional free rigid body is integrable, clearly more work is required. For instance consider the six dimensional free rigid body which has 12 dimensional adjoint orbits. Since the equations are in the form of a Lax pair $\dot{M} = [M, \Omega]$, the spectrum of $M$ is conserved; therefore, any symmetric function of the eigenvalues are conserved as well. Thus, the functions $F_k = \text{tr}(M^k)$ for $M \in \mathfrak{so}(6)$ and $k = 2, \ldots, 6$ will be conserved along adjoint orbits; however for $k$ odd, $\text{tr}(M)^k = 0$ by the skew-symmetry of the matrix $M$. This leaves 3 non-trivial constants of the motion given by the functions $F_2$, $F_4$, and $F_6$, yet the dimension of each adjoint orbit is 12 and thus more work needs to be done for higher dimensional cases.

A very useful technique often used in systems which have a Lax pair, as in (1.37), is to introduce what is called a spectral parameter into the system. Manakov
introduces such a parameter into the system (1.37) [18].

**Theorem 1.3.2** The Hamiltonian system (1.37) can be rewritten in the following form

\[
(M + J^2\lambda)\dot{\lambda} = [M + J^2\lambda, \Omega + J\lambda]
\]

(1.43)

where \( \Omega \in \mathfrak{so}(n) \), \( M = J\Omega + \Omega J \in \mathfrak{so}(n) \), \( J \) is the diagonalized matrix of the inertia map, and \( \lambda \in \mathbb{R} \) is the spectral parameter.

**Proof** The proof is a quick calculation:

\[
(M + \lambda J^2)\dot{\lambda} = [M + \lambda J^2, \Omega + \lambda J]
\]

\[
= [M, \Omega] + \lambda ([J^2, \Omega] + [M, J]) + \lambda^2 [J^2, J]
\]

\[
= [M, \Omega].
\]

The last line follows since \([J, J^2] = 0\) and \([J^2, \Omega] = [J, M]\). □

Using the equation in the form (1.43), one hopes to generate enough functions on \( \mathfrak{so}(n) \) which satisfy the conditions on integrability (1.3.1). As we previously mentioned, the functions \( F_k = \text{tr}(X^k) = \text{tr}(M + J^2\lambda)^k \) for \( k = 2, 3, \ldots, n \) are conserved quantities, and moreover, the coefficients corresponding to each \( \lambda^j \) term in the expansion of each \( F_k \) will be constants of the motion. The matrix \( M + J^2\lambda \) is no longer skew-symmetric and thus the functions \( F_k \) for \( k \) odd will no longer be trivial. So initially it appears that there are \( 3 + 4 + \ldots + (n + 1) = \frac{(n+1)(n+2)}{2} - 3 \) conserved quantities corresponding to the coefficients of \( \lambda^j \) (for \( j = 0, 1, 2, \ldots, k \)) in each expansion of \( F_k \) (where \( k = 2, 3, \ldots, n \)) which is much more than one half the dimension of the adjoint orbits in \( \mathfrak{so}(n) \).

The initial count is however a gross over estimate, as many of the coefficients are trivial or constant on a fixed adjoint orbit. For instance, the coefficients associated to the \( \lambda^0 \) term will be \( M^k \) for each \( k \) and since the restriction of \( \text{tr}(M^k) \) to an adjoint
orbit is constant, these are trivial Casimirs. Similarly, the coefficients $\text{tr}(J^{2k})$ of the $\lambda^k$ term in the expansion of $F_k$ are constants as $J$ is a fixed diagonal matrix. Lastly, note that the coefficients of each term $\lambda^{k-j}$ where $k-j$ is odd will be zero, since this coefficient consists of the sum of terms of the form:

$$\text{tr} (M^{k-j} J^j) = \text{tr} (J^j (M^{k-j})^t) = -\text{tr} (M^{k-j} J^j) = 0,$$

where $M^{k-j}$ is skew-symmetric ($k-j$ is odd) and $J^j$ is diagonal.

Now, counting the remaining non-trivial coefficients, if $k$ is even then $F_k$ will give rise to $(k+1) - 2 - \frac{k}{2} = \frac{k-2}{2}$ non-trivial constants of the motion. Likewise for $k$ odd, there will be $(k+1) - 2 - \frac{k-1}{2} = k-1$ non-trivial constants of the motion. Summing over all functions $F_k$ for $k = 2, \ldots, n$, if $n$ is even there are exactly $\frac{n(n-2)}{4}$ conserved quantities. Likewise, for $n$ odd there are exactly $\frac{(n-1)^2}{4}$ conserved quantities. In both cases, the total number of conserved quantities equals exactly one half the dimension of the adjoint orbits. The collection of such conserved coefficients are called Manakov’s integrals, which Manakov proves are in involution with each other [18].

Introducing the spectral parameter $\lambda$ into equation (1.43) is common trick in working with Lax pair systems. As demonstrated above, it is often used to find integrals of the motion. Also, note that the spectrum of the matrix $M + J^2 \lambda$ will be independent of the spectral parameter $\lambda$, and thus the eigenvalues $\mu_k$ of $M + J \lambda^2$ will be constant for all $\lambda$. Therefore, the polynomial

$$\text{Det} \left( (M + J^2 \lambda) - I \mu_k \right) = 0$$

will yield a Riemann surface of genus $g = \frac{n-2}{2}$ for even $n$ or genus $g = \frac{n-1}{2}$ if $n$ is odd. The solutions can be found by integrating Abelian functions, and we can find a explicit solutions in terms of ratios of Riemann theta functions; however, integrability does not play a role in our stability analysis, and we will not pursue
this avenue of work.

Returning to the general Dubrovin equation in (1.42), we now let $A$ and $B$ be arbitrary diagonal matrices such that $a_i \neq a_j$ and $b_i \neq b_j$ when $i \neq j$, and let $V$ be a matrix in $\mathfrak{sl}(n)$ with zeros along its diagonal. Ratiu proves the Dubrovin equation can be reduced to a Hamiltonian system on the adjoint orbits of $\mathfrak{sl}(n)$, and he generates a complete set of integrals [24]. In order to show this, we first express an arbitrary matrix in $\mathfrak{sl}(n)$ as the sum of a diagonal matrix $D$ and matrix $W$ which has zeros along its diagonal. Consider the following symmetric operator which maps $\mathfrak{sl}(n)$ into itself:

$$L(W)_{ij} = \left( \frac{b_i - b_j}{a_i - a_j} \right) w_{ij}, \quad (1.44)$$

with $L$ any arbitrary symmetric map on the diagonal entries. Setting $A = J^2$ and $B = J$, it follows that $L(M)_{ij} = \frac{1}{\lambda_i + \lambda_j} m_{ij} = \omega_{ij}$, and thus the map $L$ in (1.44) is the inverse of the inertia map.

Using the map $L$, we generalize the rigid body Hamiltonian to $\mathfrak{sl}(n)$, namely $H(M) = \frac{1}{2} (M, L(M)) = -\frac{1}{4} \text{tr} (ML(M))$ for $M \in \mathfrak{sl}(n)$. As in the case of the rigid body, the Hamiltonian vector field is given by the Lax pair $X_H(M) = [M, L(M)]$, and the dynamics in $\mathfrak{sl}(n)$ can be reduced to the adjoint orbits. Since the Hamiltonian system projects onto the adjoint orbits, it is sufficient to prove integrability on the adjoint orbits. The integrals arise by introducing a spectral parameter into the Lax pair. Therefore, they are of the general form of Manakov’s integrals and similarly in involution.
CHAPTER 2

REDUCED DYNAMICS: THE ENERGY-CASIMIR METHOD AND CLASSIFICATION OF GENERIC EQUILIBRIA

In this chapter we study the dynamics of the rigid body. Recall from Section 1.2.2 that the Hamiltonian function $H$ (see (1.29)), defined on $TSO(n)$, is left invariant. We therefore reduce the dynamics in two steps. Using body coordinates on the tangent bundle, we first translate the dynamics to $so(n)$. Next, we use Theorem 1.2.2 to find the equations of motion for the rigid body in $so(n)$, which we express in (2.6):

$$X_H(\Omega) = [J, \Omega^2].$$

The corresponding Hamiltonian function is given in Theorem 1.2.2 by equation (1.38, $H: g \rightarrow \mathbb{R}: \Omega \mapsto -\frac{1}{2}\text{tr}(J\Omega^2)$).

Lastly, we prove in Theorem 1.2.2 that flows of the Hamiltonian system $X_H$ are contained in adjoint orbits of $so(n)$, where we reduce the dynamics further.

We handle the even and odd cases of the rigid body separately, since the underlying geometries of the Lie algebras $so(2n+1)$ and $so(2n)$ are fundamentally different. In particular, the dimension of the non-degenerate adjoint orbits of $so(2n+1)$ are $\text{dim}(so(2n+1)) - \text{rank}(so(2n+1)) = 2n^2$, while the dimension of the non-degenerate adjoint orbits of $so(2n)$ are $\text{dim}(so(2n)) - \text{rank}(so(2n)) = 2n^2 - 2n$. The dynamics of the rigid body for the even and odd dimension rigid bodies must therefore be treated as separate cases. Now we will focus our attention solely on the dynamics for an arbitrary even dimensional rigid body, which has configuration space $SO(2n)$ and Lie algebra $so(2n)$. We expect to handle the odd dimensional case in similar
fashion as in Chapter 3.

In Section 2.1, we define the non-degenerate adjoint orbits in $\mathfrak{so}(2n)$ as the level sets of $n$ linearly independent Casimirs of the Hamiltonian system $X_H$. Next, in Section 2.2 we introduce a general notion of equilibrium for reduced Hamiltonian systems. In Section 2.3 we discuss stability and introduce the energy-Casimir method. Lastly, in Section 2.4 we prove a corollary to Theorem 2.4.6, which classifies all generic equilibria of the general $2n$ dimensional rigid body. The classification of the generic equilibria is proved by Fehér and Marshall [10].

2.1 Structure of Non-Degenerate Adjoint Orbits

First we recall the identification of the Lie algebra $\mathfrak{so}(2n)$ with its dual $\mathfrak{so}^*(2n)$ induced by the form $(X, Y) = -\frac{1}{2} \text{tr}(XY)$ (see Section 1.2.1). Thus, the Lie algebra $\mathfrak{g} = \mathfrak{so}(2n)$ can be equipped with a Poisson structure induced by the Lie-Poisson structure on $\mathfrak{so}^*(2n)$. Namely, for any two smooth, real-valued functions $F$ and $G$ on $\mathfrak{g}$, we define the Poisson bracket on $\mathfrak{so}(2n)$ as

$$\{ F, G \}(\Omega) = (\Omega, [dF(\Omega), dG(\Omega)]) = -\frac{1}{2} \text{tr} (\Omega \, [dF(\Omega), dG(\Omega)]) .$$

(2.1)

Since we are justified in identifying linear functionals, such as $dF$ and $dG$, with skew-symmetric matrices in $\mathfrak{so}(2n)$, the Poisson bracket in (2.1) makes sense.

Using the same notation as in Chapter 1, we denote the adjoint orbit through a $\Omega \in \mathfrak{so}(2n)$ by

$$\sigma_\Omega = \{ \text{Ad}_g \Omega \mid g \in SO(2n) \} .$$

Throughout our stability analysis of the even dimensional rigid body, we assume $\sigma_\Omega$ is a generic adjoint orbit with dimension $\dim (\mathfrak{g}_\Omega) = \dim (\mathfrak{so}(2n)) - \text{rank}(\mathfrak{so}(2n)) = 2n^2 - 2n$. We now consider motion of the $2n$ dimensional rigid body, whose motion is governed by the Hamiltonian system obtained in Theorem 1.2.2. As a result of the theorem, the system $X_H = [M, \Omega]$ projects to a Hamiltonian system on the
adjoint orbits, which are submanifolds of the quotient space defined by the Casimir functions. Therefore, we define the adjoint orbits as the level sets of \( n \) linearly independent Casimirs.

Euler’s equation for the rigid body is expressed by the Lax pair in equation (1.37), and therefore the spectrum of \( \Omega(t) \) is constant. Moreover, any symmetric function of the eigenvalues will be constant as well, and therefore any function of the form \( F_k = \text{tr}(\Omega^k) \) is constant on a fixed adjoint orbit. Since we are working with skew-symmetric matrices in \( \mathfrak{so}(2n) \), \( F_k(\Omega) = 0 \) when \( k \) is odd. Using the characteristic polynomial for a given \( \Omega \in \mathfrak{so}(2n) \), we can express \( F_{2n}(\Omega) = \text{tr}(\Omega^{2n}) \) in terms of the trace functions \( F_2, F_4, \ldots, F_{2n-2} \). Instead we introduce another, linearly independent, symmetric function of the eigenvalues called the Pfaffian, \( Pf(\Omega) = \sqrt{\det(\Omega)} \), and from now on we consider the following collection of \( n \) linearly independent Casimirs:

\[
C_1(\Omega) = -\frac{1}{4}\text{tr}(\Omega^2)
C_2(\Omega) = \frac{1}{8}\text{tr}(\Omega^4)
\vdots
C_k(\Omega) = \frac{(-1)^k}{4k}\text{tr}(\Omega^{2k})
\vdots
C_{n-1}(\Omega) = \frac{(-1)^{2n-1}}{4(n-1)}\text{tr}(\Omega^{2n-2})
C_n = Pf(\Omega) = \sqrt{\det(\Omega)}.
\]

The leading factors will be very useful in simplifying calculations in Chapter 3.

2.2 Relative Equilibria

The concepts of equilibrium and stability change depending on which stage of reduction we are in. We now state exactly what type of dynamics these terms refer to in the different stages of reduction. We now consider a general mechanical system with symmetry on an unreduced phase space \( P = T^*G \) (for \( G \) a Lie group). Let the symmetry group \( G \) act on itself, and let \( H \) be the \( G \)-invariant Hamiltonian
on defined on $T^*G$. We can perform the coadjoint orbit reduction, and study the reduced dynamics in $\mathfrak{g}^*$ (now being careful to distinguish between $\mathfrak{g}$ and $\mathfrak{g}^*$).

Due to the symmetry of the system on the unreduced phase space, the simplest notion of equilibrium refers to points of phase space which lie on flows which are one parameter orbits of the symmetry group $G$. Such points are called relative 
*equilibria*, and the solutions of through these points flow only in the direction of the symmetry of the system. Since the solution starting at a relative equilibrium $p_e \in T^*G$ is thus a group orbit, in the reduced Poisson system, it projects to a stationary equilibrium $\pi(p_e) \in \mathfrak{g}^*$. We will be concerned only in determining the stability of relative equilibria which project onto non-degenerate adjoint orbits. Such relative equilibria are called *regular*. Since relative equilibria on the unreduced phase space are synonymous with fixed point equilibria in the reduced space, the term equilibria will refer to both types of solutions in $T^*G$ and $\mathfrak{g}^*$.

### 2.3 The Energy-Casimir Method

Next, we clarify our notion of stability of these equilibria in each setting (both reduced and unreduced). Lyapunov’s classical method provides sufficient conditions for Lyapunov stability of a stationary equilibrium, $x_e$, of a conservative system. In general, the hope is to find a conserved quantity called a *Lyapunov function* (normally the energy) for which $x_e$ is a local maximum or minimum, in which case solutions that start near the stationary solution will remain close to the equilibrium. If $\mu_e \in \mathfrak{g}^*$ is a local extremum of the restriction of the conserved Hamiltonian to the coadjoint orbit which contains $\mu_e$, then Lyapunov’s classical method suggests $\mu_e$ is a stable fixed point solution of the flow on the coadjoint orbit. We call such an equilibrium *leafwise stable*. It does not necessarily follow that a leafwise stable equilibrium is Lyapunov stable on the Poisson system (in $\mathfrak{g}^*$), since $\mu_e$ might not be a critical point of the Hamiltonian $\hat{H}$ (on all of $\mathfrak{g}^*$, not restricted to the coadjoint orbits).
In a major development in the study of reduced dynamics on the coadjoint orbits of Lie-Poisson systems, Arnold extends Lyapunov’s classical method to study the behavior of such equilibrium which are not critical points of the energy [2]. Arnold applies his method to analyze the stability of planar ideal incompressible fluid motion. Arnold’s method, now called the energy-Casimir method, uses the symmetry of the system to generate a Casimir function $C$, which is conserved along the coadjoint orbits. This Casimir (if it exists) is added to the Hamiltonian $\tilde{H}$ in the hopes of creating a new, suitable Lyapunov function. Arnold’s method works provided the equilibrium is a regular point in $g^*$, in which case all of the coadjoint orbits near the equilibrium have the same dimension. Arnold’s method is later developed by Holm, Marsden, Ratiu, and Weinstein, who have applied the method to various systems, including the three dimensional rigid body [12].

**Theorem 2.3.1** Let $X_H$ be a Hamiltonian system on $T^*G$ which is $G$-invariant. Let $\tilde{X}_H$ be the reduction of the system to $g^*$ with reduced energy $\tilde{H} : g^* \to \mathbb{R}$, and let $\mu_e$ be a regular point in $g^*$ which is an equilibrium of the reduced system. If there exists a Casimir function $C : g^* \to \mathbb{R}$ such that

1. $d(\tilde{H} + C)(\mu_e) = 0$
2. $d^2(\tilde{H} + C)(\mu_e)$ is positive or negative definite,

then $\mu_e$ is a Lyapunov stable equilibrium of $\tilde{X}_H$.

Figure (2.3) illustrates how the method uses the symmetry of the system to create a Lyapunov function. The planes represent the symplectic leaves, and since $\mu_e$ is regular, they all have the same dimension. The parabolas represent the level sets of the conserved function $\tilde{H} + C$. Since both the symplectic leaves and the function $\tilde{H} + C$ are constant on along the flows of the system, the flows must lie in the ellipsoid-like hypersurfaces of their intersection. Since $\mu_e$ is a critical point, it lies in a fixed point orbit. Notice that a flow through a point near $\mu_e$, which lies
on the same leaf as $\mu_e$, is an ellipsoid-like orbit which will remain close to $\mu_e$. Since $\mu_e$ is a regular point and $\tilde{H} + C$ is definite, in every leaf near $\mu_e$ the level sets of $\tilde{H} + C$ form a family of ellipsoid-like hypersurfaces, which are centered around an equilibrium. Therefore, any point near $\mu_e$ (either on the same leaf or on another leaf near $\mu_e$) lies on a solution of $\tilde{X}_H$ which lies inside of an ellipsoid. Thus $\mu_e$ is Lyapunov stable.

Figure 2.1: An illustration of Arnold’s energy-Casimir method. Flows in the intersections of the symplectic leaves and level sets of modified Hamiltonian $\tilde{H} + C$ lie inside ellipsoid-like hypersurfaces, and remain in a neighborhood near the equilibrium $\mu_e$. [4]

The energy-Casimir method provides a systematic approach to determining stability of equilibria in reduced Lie-Poisson systems which are regular points. It is important to comment on the limitations of this method. As in the case of Lyapunov’s method, the energy-Casimir method only provides sufficient conditions for an equilibrium to be stable. If the Casimir is indefinite at an equilibrium, the method is inconclusive, as the equilibrium still might be stable. Arnold developed his method in order to study the dynamics near equilibria of systems in fluid dynamics, which are reduced systems on Lie-Poisson algebras. However, the energy-Casimir method
applies only to Arnold’s Lie group setting, and it cannot be applied to a large array of mechanical systems with symmetry whose configuration is a general manifold (not necessarily a Lie group). Other methods have been recently developed to handle these more general cases, most notably the energy-Momentum method [27], which is first formalized by Smale [28].

Before moving on to the particular setting of the rigid body, we lastly comment on how stability in the reduced Poisson manifold carry over into the unreduced phase space $T^*G$. For example, consider a three dimensional rigid body which is rotating in the direction of the longest or shortest principal axis of inertia. By rotating the body faster, but in the same direction, we jump to a nearby symplectic leaf in the reduced space, which stays close to the equilibrium. However, the two rotations yield two distinct paths in the unreduced phase space, and we wish to study the distance between the two orbits in the $T^*G$. At any point in time, the endpoints of each might be far (with respect to some topology on $T^*G$) from each other, but it is possible to bring the two flows back together by multiplying by some element of the symmetry group [25]. We call such a relative equilibrium $G$-stable:

**Theorem 2.3.2**  Let $P$ denote the unreduced phase space of a system with symmetry group $G$. If $G$ acts freely and properly on $P$, then a relative equilibrium $p_e \in P$ is $G$-stable if the projection onto the stationary equilibrium $\mu_e = \pi(p_e) \in P/G$ is Lyapunov stable with respect to the reduced dynamics in $P/G$.

2.4 Generic equilibria

We now return to the example of the $2n$ dimensional rigid body. We wish to study the dynamics of relative equilibria in the phase space, and since we are in the Lie-Poisson setting of the Lie algebra $\mathfrak{so}(2n)$, relative equilibria are stationary equilibria in the reduced Lie-Poisson dynamics on $\mathfrak{g}$. We restrict our attention to relative equilibria which are regular points in the unreduced phase space, meaning they lie
on non-degenerate adjoint orbits in the reduced Lie-Poisson setting in $\mathfrak{g}$. Recall from Section 2.1 that such orbits are level sets of $n$ linearly independent Casimir functions. The non-degenerate orbits have dimension $2n^2 - 2n$ (codimension equal to the rank of the Lie algebra $\mathfrak{so}(2n)$). Secondly, from now on we assume the inertia matrix $J$ we found in 1.36 is a regular diagonal matrix:

$$J = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{2n})$$

such that for $\lambda_i \neq \lambda_j$ when $i \neq j$. \hspace{1cm} (2.3)

We call equilibria for a rigid body with a regular inertia matrix, which lie on non-degenerate adjoint orbits, \textit{generic equilibria}. For the rest of our discussion, we study the stability of stationary equilibria of the reduced Hamiltonian system on $\mathfrak{g}$ (see Theorem 1.2.2). Stability of these stationary, generic equilibria imply relative equilibria are $G$-stable in the unreduced phase space (see Section 2.3). Next, as a result of the work of Fehér and Marshall, we classify all generic equilibria on a given non-degenerate adjoint orbit \cite{10}.

Let $\mathfrak{h}$ denote the Cartan subalgebra of $\mathfrak{so}(2n)$ spanned by matrices of the following form:

$$Y_d^j = e_{jj} \otimes i\sigma_2 = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & -1 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \quad \text{for } 1 \leq j \leq n$$ \hspace{1cm} (2.4)

where $e_{jj}$ is the $n \times n$ matrix with 1 in the $j^{\text{th}}$ diagonal entry and zeros elsewhere,
and where $\sigma_2$ denotes the usual Pauli matrix $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$. The matrices in (2.4) span a maximal torus in $\mathfrak{so}(2n)$, and we will refer to these matrices in the maximal torus as quasi-diagonal.

A generic element in the Cartan subalgebra $\mathfrak{h}$ can be expressed via the matrices $Y^j_d$,

$$h = \sum_{j=1}^n a_j Y^j_d = \begin{pmatrix} 0 & a_1 & 0 & 0 & \cdots & 0 & 0 \\
-a_1 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & a_2 & \cdots & 0 & 0 \\
0 & 0 & -a_2 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & 0 & \cdots & 0 & a_n \\
0 & 0 & 0 & 0 & \cdots & -a_n & 0 \end{pmatrix}$$

for real numbers $a_j$ such that $a_i \neq a_j$ if $i \neq j$. It will be convenient to denote generic adjoint orbits by $\sigma_h$ since each generic orbit $\sigma_\Omega$ contains regular elements in the Cartan subalgebra such as $h$ in (2.5).

**Lemma 2.4.1** An element $Y \in \sigma_h$ is an equilibrium of the Hamiltonian system (1.38) if and only if $Y^2$ is a diagonal matrix.

**Proof** For any $Y \in \sigma_h$, the equations of motion for the rigid body given in (1.37) can be expressed as

$$X_H(Y) = [M, Y] = (JY + YJ)Y - Y(JY + YJ) = JY^2 - Y^2J = [J, Y^2].$$

Since we have assumed in (2.3) that $J$ is a regular diagonal matrix, it immediately
follows that $J$ and $Y^2$ commute if and only if $Y^2$ is a diagonal matrix.

It follows immediately from Lemma 2.4.1 that any element $h \in \mathfrak{h}$ is a generic equilibrium since

$$
\begin{pmatrix}
-a_1^2 & 0 & \cdots & 0 \\
0 & -a_1^2 & \cdots & \vdots \\
& -a_2^2 & \ddots & \vdots \\
& & \ddots & -a_2^2 \\
& & & \vdots \\
0 & \cdots & & -a_n^2
\end{pmatrix}
$$

However, matrices $h \in \mathfrak{h}$ do not exhaust all such equilibria. We will now classify all generic equilibria of the arbitrary, even dimensional rigid body. Fehér and Marshall classify generic equilibria for any general rigid body (both even and odd) [10]. We will classify the generic equilibria in the spirit of Fehér and Marshall’s original proof. First, we will consider an action of the permutation group $S_{2n}$ on the collection of all diagonal $2n \times 2n$ matrices denoted $\text{Diag}(2n)$. Then we will prove several rather basic lemmas regarding this action which will be essential in the proof of the theorem to follow. To begin, we introduce the set

$$
E_h = \{ Y = ghg^{-1} \in \sigma_h \mid [J, gh^2g^{-1}] = 0 \ , \ g \in O(2n) \} \, .
$$

(2.7)

It is important to note the set $E_h$ is not the set of all equilibria on $\sigma_h$ since we have allowed conjugation by $g \in O(2n)$ instead of restricting to $g \in SO(2n)$. We will account for this discrepancy shortly.

Let $p$ denote an element in the permutation group $S_{2n}$. Identify each element
\[ p \in S_{2n} \text{ with a permutation matrix } \bar{p} \text{ in } O(2n) \text{ via the map} \]
\[ \psi : S_{2n} \rightarrow O(2n) : p \mapsto \bar{p} \quad \text{where } \bar{p}_{ij} = \delta_{i, \bar{p}(j)} \]

The identification \( \psi \) induces an action of the group \( S_{2n} \) on \( \text{Diag}(2n) \) in the following way:
\[ S_{2n} \times \text{Diag}(2n) \rightarrow \text{Diag}(2n) : p(D) \mapsto \bar{p}D\bar{p}^{-1}. \quad (2.8) \]

Conjugating the equilibrium \( h \) by a permutation matrix \( \bar{p} \) will thus yield a matrix \( Y = \bar{p}hp^{-1} \) whose square \( Y^2 = \bar{p}hp^{-1} \) is the diagonal matrix obtained by the corresponding permutations of the diagonal entries of \( h^2 \).

**Lemma 2.4.2** Let \( \tau_{i,i+1} \in S_{2n} \) denote a transposition of consecutive terms and consider the following subgroup of \( S_{2n} \),
\[ T = \langle \tau_{1,2}, \tau_{3,4}, \ldots, \tau_{2n-1,2n} \rangle, \quad (2.9) \]

generated by transpositions of the form \( \tau_{2i-1,2i} \) for \( 1 \leq i \leq n \). A permutation \( p \in S_{2n} \) fixes the diagonal matrix \( h^2 \) with respect to the action \( p(h^2) = \bar{p}h^2\bar{p}^{-1} \) if and only if \( p \in T \).

**Proof** The action of a transposition \( \tau_{2i-1,2i} \in S_{2n} \) on \( h^2 \) will permute the diagonal entries in rows \( 2i - 1 \) and \( 2i \) of \( h^2 \) which are both \( -a_i^2 \). Thus any composition of such transpositions will leave the diagonal entries of the matrix \( h^2 \) unchanged.  

**Lemma 2.4.3** Consider the set of block diagonal matrices in \( O(2n) \) with blocks of \( 2 \times 2 \) orthogonal matrices in \( O(2) \)
\[ \left\{ r \in O(2n) \left| r = \sum_{j=1}^{n} (e_{jj} \otimes R_j) \ , \ R_j \in O(2) \right\} \right. , \]

\[ \]
where $e_{jj}$ is as described in (2.4). A matrix $r$ in $O(2n)$ is in the isotropy group $O(2n)_{h^2}$ of $h^2$ if and only if $r$ is an element of this set.

Proof The matrix $h^2$ can be viewed as a block diagonal matrix with $n$ distinct $2 \times 2$ blocks (since $a_i \neq a_j$) of the form

$$
\begin{pmatrix}
-a_j^2 & 0 \\
0 & -a_j^2
\end{pmatrix}
$$

for $1 \leq j \leq n$. Generalizing a basic result about regular diagonal matrices to our current $2 \times 2$ block diagonal setting: the only matrices which commute with a $2n \times 2n$ block diagonal matrix with $n$ distinct $2 \times 2$ blocks such as $h^2$ are $2n \times 2n$ block diagonal matrices with $2 \times 2$ blocks along the diagonal. Since we are only concerned about orthogonal matrices which fix $h^2$, we can restrict ourselves to matrices in $O(2n)$ with $n$ arbitrary blocks $R_j \in O(2)$ along the diagonal. It is routine to check that all such matrices $r$ in fact satisfy $rh^2r^{-1} = h^2$. The isotropy group of $h^2$ is therefore

$$
O(2n)_{h^2} := \{ r \in O(2n) \mid rh^2r^{-1} = h^2 \} = \left\{ r \in O(2n) \left| r = \sum_{j=1}^{n} (e_{jj} \otimes R_j) , R_j \in O(2) \right. \right\}. \tag{2.10}
$$

Lemma 2.4.4 Consider the set of block diagonal matrices in $O(2n)$ with blocks of $2 \times 2$ matrices in $SO(2)$

$$
\left\{ q \in O(2n) \left| q = \sum_{j=1}^{n} (e_{jj} \otimes Q_j) , Q_j \in SO(2) \right. \right\}.
$$

A matrix $q$ in $O(2n)$ is in the isotropy group $O(2n)_h$ of $h$ if and only if $q$ is a matrix of this set.

Proof We begin with the same observation as in the previous proof, namely any matrix in $O(2n)_h$ must likewise be block diagonal with $n$ blocks of size $2 \times 2$ along its diagonal. Again we must restrict our attention to such matrices in $O(2n)$. Since each $2 \times 2$ block of a potential stabilizer will transform only one block of $h$ and leave
the others unchanged, it is sufficient to prove for each individual \(1 \leq j \leq n\) that

\[
Q_j \begin{pmatrix} 0 & a_j \\ -a_j & 0 \end{pmatrix} Q_j^{-1} = \begin{pmatrix} 0 & a_j \\ -a_j & 0 \end{pmatrix}
\]

if and only if \(Q_j \in SO(2)\). This result can then be extended simultaneously to all blocks of \(h\). Consider an arbitrary matrix \(Q_j = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} \in O(2)\) and compute the product

\[
Q_j \begin{pmatrix} 0 & a_j \\ -a_j & 0 \end{pmatrix} Q_j^{-1} = Q_j \begin{pmatrix} 0 & a_j \\ -a_j & 0 \end{pmatrix} Q_j^t
\]

\[
= \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} \begin{pmatrix} 0 & a_j \\ -a_j & 0 \end{pmatrix} \begin{pmatrix} q_{11} & q_{21} \\ q_{12} & q_{22} \end{pmatrix}
\]

\[
= \begin{pmatrix} 0 & a_1(q_{11}q_{22} - q_{12}q_{21}) \\ -a_1(q_{11}q_{22} - q_{12}q_{21}) & 0 \end{pmatrix}.
\]

Clearly \(Q_j\) fixes each \(2 \times 2\) block if and only if \(\det(Q_j) = 1\) for all \(1 \leq j \leq n\). Therefore the isotropy group of \(h\) can be expressed as follows:

\[
O(2n)_h := \{ q \in O(2n) \mid qhq^{-1} = h \}
\]

\[
= \left\{ q \in O(2n) \mid q = \sum_{j=1}^{n} (e_{jj} \otimes Q_j) , Q_j \in SO(2) \right\}.
\]

(2.11)

**Lemma 2.4.5** Let \(T\) denote the subgroup of \(O(2n)\) generated by matrices \(\tau_{2i-1,2i}\) corresponding to each \(\tau_{2i-1,2i} \in T\). Each \(r \in O(2n)_{h^2}\) can be written uniquely as the product \(r = \bar{\tau}q\) for a \(\bar{\tau} \in \overline{T}\) and \(q \in O(2n)_h\).

**Proof** The group \(SO(2)\) is a normal subgroup of \(O(2)\) and the corresponding quo-
tient group is
\[ T = \left\{ I_2, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\} = \{ id, \bar{\tau}_{1,2} \}, \]
and thus it follows that any matrix \( R \in O(2) \) can be written uniquely as the product
\[ R = \bar{\tau}Q \]
for a \( Q \in SO(2) \) and \( \bar{\tau} \in T \). In particular if \( R \in SO(2) \) then \( \bar{\tau} = id \) and if
\( R \notin SO(2) \) then \( \bar{\tau} = \bar{\tau}_{1,2} \).

Now consider the isotropy group \( O(2n)_{h^2} \) with arbitrary \( 2 \times 2 \) blocks of matrices
\( R_j \in O(2) \) for \( 1 \leq j \leq n \). We can decompose each block \( R_j \) uniquely into a product
\[ R_j = \bar{\tau}^j Q_j \]
where \( Q_j \in SO(2) \) and \( \bar{\tau}^j = I_2 \) or \( \bar{\tau}^j = \bar{\tau}_{2j-1,2j} \). This decomposition can be
performed simultaneously for any \( r \in O(2n)_{h^2} \):

\[
\begin{pmatrix}
R_1 & 0 & \ldots & 0 \\
0 & R_2 & \ldots & \\
\vdots & \vdots & \ddots & \\
0 & \ldots & 0 & R_n
\end{pmatrix} = \begin{pmatrix}
\bar{\tau}^1 Q_1 & 0 & \ldots & 0 \\
0 & \bar{\tau}^2 Q_2 & \ldots & \\
\vdots & \vdots & \ddots & \\
0 & \ldots & 0 & \bar{\tau}^n Q_n
\end{pmatrix} = \bar{\tau}q
\]
with \( \bar{\tau} = \prod_{j=1}^n \bar{\tau}^j \in T \) and \( q = \sum_{j=1}^n (e_{jj} \otimes Q_j) \in O(2n)_h \).

Using the lemmas above we now proceed to the statement and proof the main
theorem of this section due to Fehér and Marshall [10]. After completing the proof
of the theorem we will account for the slight discrepancy between the definition of
the set \( E_h \) in (2.7) and the actual set of generic equilibria which must lie in \( \sigma_h \).

**Theorem 2.4.6** [10] Let \( h \) be a generic element in the Cartan subalgebra \( \mathfrak{h} \) as in
(2.5). A matrix \( Y = ghg^{-1} \) for \( g \in O(2n) \) is contained in the set \( E_h \) defined by equation (2.7) if and only if there exists a unique \( p \in S_{2n} \) such that \( Y = \bar{p}h\bar{p}^{-1} \).

**Proof** For any \( p \in S_{2n} \) and \( Y = \bar{p}h\bar{p}^{-1} \) it is clear that \( Y^2 = \bar{p}h^2\bar{p}^{-1} = p(h^2) \) is in fact a diagonal matrix since the action of \( p \) on \( h^2 \) merely permutes the diagonal entries of \( h^2 \). Indeed any \( Y = \bar{p}h\bar{p}^{-1} \) is an element in \( E_h \) since it is diagonal. Next we show if \( Y = ghg^{-1} \) is in \( E_h \), then there exists a unique \( p \in S_{2n} \) such that \( Y = \bar{p}h\bar{p}^{-1} \).

Assume \( Y = ghg^{-1} \in E_h \). Since \( Y^2 = gh^2g^{-1} \) is a diagonal matrix obtained by conjugating \( h^2 \) by an orthogonal matrix \( g \) from \( O(2n) \), \( Y^2 \) must be a diagonal matrix whose entries appear as permutations of the original entries of \( h^2 \). Consider the quotient of \( S_{2n} \) by \( T \), the isotropy subgroup of \( h^2 \) defined in (2.9),

\[
S_{2n}/T = \{p_i\}_{i=1}^N. \tag{2.12}
\]

In particular we know \( \{p_i\}_{i=1}^N \) contains \( N = \frac{(2n)!}{2^n} \) permutations such that \( p_i(h^2) \neq p_j(h^2) \) if \( i \neq j \), and all possible permutations of \( h^2 \) are exhausted by the elements of \( \{p_i\}_{i=1}^N \). Thus if \( Y = ghg^{-1} \in E_h \) we know there exists a unique permutation \( p_i \) from (2.12) such that \( Y^2 = gh^2g^{-1} = \bar{p}_ih^2\bar{p}_i^{-1} \). We have equivalently shown

\[
(\bar{p}_i^{-1}g)h^2(\bar{p}_i^{-1}g)^{-1} = h^2.
\]

For any \( g \in O(2n) \) such that \( ghg^{-1} \in E_h \), by Lemma 2.4.3 we know \( \bar{p}_i^{-1}g = r \) for an \( r \in O(2n)_{h^2} \), and for any general such \( g \) we know

\[
g = \bar{p}_ir \tag{2.13}
\]

for a unique \( p_i \in S_{2n}/T \) and \( r \in O(2n)_{h^2} \).

Recall from Lemma 2.4.5 that each \( r \in O(2n)_{h^2} \) can be written uniquely as a product \( r = \bar{\tau}q \) for a matrix \( \bar{\tau} \) corresponding to a product of transpositions \( \tau \in T \) and a \( q \in O(2n)_h \). Thus in equation (2.13) we can further conclude that if
\[ Y = ghg^{-1} \in E_h, \text{ then it must follow that } g \in O(2n) \text{ can be written as the unique product of matrices} \]
\[ g = \tilde{p}_i \tilde{\tau} q. \quad (2.14) \]

Finally we have shown that if \( Y \in E_h \), it follows
\[
Y = ghg^{-1} \\
= (\tilde{p}_i \tilde{\tau} q) h (\tilde{p}_i \tilde{\tau} q)^{-1} \\
= (\tilde{p}_i \tilde{\tau}) (ghq^{-1}) (\tilde{p}_i \tilde{\tau})^{-1} \\
= \tilde{p}h\tilde{p}^{-1}
\]
since \( q \in O(2n)_h \). Moreover the corresponding permutation \( p = p_i \tau \) is unique since both \( \tilde{p}_i \) in (2.12) and \( \tilde{\tau} \) obtained in Lemma 2.4.5 are unique. Therefore if \( Y = ghg^{-1} \) is an element from \( E_h \), there exists a unique \( p \in S_{2n} \) such that \( Y = \tilde{p}h\tilde{p}^{-1} \). Indeed different permutations \( p \) give rise to distinct elements \( Y \) in \( E_h \). \( \blacksquare \)

Theorem 2.4.6 classifies all \( g \in O(2n) \) such that \( Y \in E_h \); however we only wish to find generic equilibria contained in adjoint orbits such as \( \sigma_h \). We now restrict our attention to \( Y \in \sigma_h \) which belong to the set of all generic equilibria:

\[
\tilde{E}_h := \{ Y = ghg^{-1} \mid [J, Y^2] = 0 \text{ for } g \in SO(2n) \}.
\quad (2.15)
\]

The discrepancy in the definitions of the sets \( E_h \) and \( \tilde{E}_h \) can be accounted for in the next lemma, which in conjunction with Theorem 2.4.6, will give a complete classification of all generic equilibria.

**Lemma 2.4.7** Let \( A_{2n} \) denote the subgroup of \( S_{2n} \) of even permutations. A permutation matrix \( \tilde{p} \) is an element of \( SO(2n) \) if and only its corresponding permutation \( p \) is an even permutation in \( A_{2n} \).

**Proof** If \( p \in A_{2n} \), the corresponding permutation matrix \( \tilde{p} \) is obtained by perform-
ing an even number of row permutations to the identity matrix. Conversely if \( p \) is an odd permutation, \( \bar{p} \) is result of performing an odd number of row permutations to the identity matrix. Thus it follows that \( \text{sgn}(p) = \det(\bar{p}) \) and \( \bar{p} \in SO(2n) \) if and only if \( \text{sgn}(p) = 1. \]

Lemma 2.4.7 can be applied to the result of Theorem 2.4.6 to obtain the corollary below,

**Corollary 2.4.8** Let \( h \) be a generic element in the Cartan subalgebra \( \mathfrak{h} \) in (2.5). For \( g \in SO(2n) \), a matrix \( Y = ghg^{-1} \) is contained in the set \( \tilde{E}_h \) (2.15) if and only if there exists a unique even permutation \( p \in A_{2n} \) such that \( Y = \bar{p}h\bar{p}^{-1} \).

Suppose now wish to study the behavior of a generic equilibrium \( \bar{p}h\bar{p}^{-1} \in \sigma_h \) of a given rigid body with corresponding inertia matrix \( J = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{2n}) \). Using the equations of motion for the rigid body derived in (1.37), we study the behavior of the system,

\[
\frac{d}{dt}(\bar{p}h\bar{p}^{-1}) = [J, \bar{p}h\bar{p}^{-1}] = J\bar{p}h\bar{p}^{-1} - \bar{p}h\bar{p}^{-1}J. \tag{2.16}
\]

Equivalently we can apply a change of basis, according to the permutation matrix \( \bar{p} \), to the system such that the corresponding inertia matrix becomes \( \tilde{J} = \bar{p}^{-1}J\bar{p} \). Rewriting the equations of motion, (2.16) can now be expressed with respect to the new basis as follows:

\[
\frac{d}{dt}(\bar{p}h\bar{p}^{-1}) = J\bar{p}h\bar{p}^{-1} - \bar{p}h\bar{p}^{-1}J
= \bar{p}\tilde{J}\bar{p}^{-1} - \bar{p}h\tilde{J}\bar{p}^{-1}
= \bar{p} \left[ \tilde{J}, h \right] \bar{p}^{-1}. \tag{2.17}
\]

The two equations (2.16) and (2.17) imply \( \bar{p} \left[ \tilde{J}, h \right] \bar{p}^{-1} = \frac{d}{dt}(\bar{p}h\bar{p}^{-1}) = \bar{p}h\bar{p}^{-1} \), and therefore

\[
\dot{h} = \left[ \tilde{J}, h \right] \quad \text{is equivalent to} \quad \frac{d}{dt}(\bar{p}h\bar{p}^{-1}) = [J, \bar{p}h\bar{p}^{-1}] . \tag{2.18}
\]
Summarizing the work above, given any generic equilibrium $\bar{p} h \bar{p}^{-1}$ for the rigid body with inertia matrix $J$, we permute the diagonal entries of the original matrix $J$ via the conjugation $\bar{J} = \bar{p}^{-1} J \bar{p}$. The behavior of the equilibrium $\bar{p} h \bar{p}^{-1}$ with respect to the matrix $J$ is equivalent to the behavior of the equilibrium $h$ of the form (2.5) with respect to the matrix $\bar{J}$. Thus without any loss of generality, we can assume a generic equilibrium is a regular element $h$ in the Cartan subalgebra $\mathfrak{h}$. Furthermore, we can perform a permutation such that the matrix $h$ in the Cartan subalgebra is of the special form

$$
\Omega_e = \begin{pmatrix}
0 & a_1 & 0 & 0 & \ldots & 0 & 0 \\
-a_1 & 0 & 0 & 0 & \ldots & 0 & 0 \\
0 & 0 & a_2 & 0 & \ldots & 0 & 0 \\
0 & 0 & -a_2 & 0 & \ldots & 0 & 0 \\
\ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & 0 & \ldots & a_n & 0 \\
0 & 0 & 0 & 0 & \ldots & -a_n & 0
\end{pmatrix}, \text{ with } |a_1| < \ldots < |a_n|. \quad (2.19)
$$

From now on, the term generic equilibrium implies we have permuted the inertia matrix and equilibrium such that $\Omega_e = ghg^{-1}$ is written in the form (2.19).
CHAPTER 3

THE ENERGY-CASIMIR METHOD AND STABILITY OF GENERIC
EQUILIBRIA OF THE 2N DIMENSIONAL RIGID BODY

In this chapter we apply the energy-Casimir method outlined in Section 2.3. Recall
the purpose of the method: construct a Lyapunov function from a Hamiltonian
$H$ (which is not a suitable Lyapunov function) by adding an arbitrary family of
Casimirs $C$ to the Hamiltonian. If it is possible to construct a Casimir such that an
equilibrium is a critical point for the modified Hamiltonian $H + C$ (which still has
the same dynamics as the original $H$) and the second variation of the Hamiltonian
$H + C$ is positive or negative definite at the equilibrium, then the equilibrium is
Lyapunov stable.

In Section 3.1, we will first introduce an arbitrary Casimir we call $\phi_F(\Omega) =
\phi(F(\Omega))$ (see (3.4)) and a modified Hamiltonian $H_\phi = H + \phi_F$. Then in Section
3.2, we find necessary and sufficient conditions on the first partial derivatives of the
smooth function $\phi$ such that a generic equilibrium of the form (2.19) is a critical
point of the first variation $d^2H_\phi(\Omega_e)$ (see Theorem 3.2.4). Lastly in Section 3.3,
we determine conditions on the $\lambda_i$’s of the inertia matrix in (1.36) and the second
partial derivatives of $\phi$ which ensure the second variation $d^2H_\phi$ is definite at $\Omega_e$.

Determining sufficient conditions for the definiteness of $d^2H_\phi(\Omega_e)$ is quite cum-
bersons. We will approach this problem by first choosing a convenient ba-
sis for $\mathfrak{so}(2n)$ in Section 3.3.1 in order to express the second variation as a
$(2n^2 - n) \times (2n^2 - n)$ symmetric, block diagonal matrix, which we state in Corollary
3.3.9:
Finding such an expression for the second variation is especially convenient since definiteness of the entire second variation can then be analyzed by studying the definiteness of the individual blocks, which in this case are all size $2 \times 2$, except for the first $n \times n$ block $V_{n \times n} + S_{n \times n}$. At first glance, the individual entries of the matrix $d^2 H_\phi(\Omega_e)$ look hopelessly complicated. However, in Section 3.4.1 we miraculously simplify the traces and determinants of the $2 \times 2$ blocks, and we ultimately find necessary and sufficient conditions on the entries of the inertia matrix $J$ which simultaneously ensure the definiteness of all $2 \times 2$ blocks. The conditions (stated in Theorems 3.4.8 and 3.4.9 for positive and negative definiteness respectively) are analogous to the classical longest/shortest conditions for the three dimensional rigid body.

In Section 3.4.2, we find further conditions on the arbitrary function $\phi$ which ensure the first $n \times n$ block, $V_{n \times n} + S_{n \times n}$, is definite. We do not introduce any clever tricks to simplify this larger block. Instead, we find an explicit equation for $\phi$ in Lemma 3.4.10,

$$\phi : \mathbb{R}^n \to \mathbb{R} : \mathbf{x} = (x_1, \ldots, x_n) \mapsto \mathbf{x} \cdot \Phi' + c \left( \sum_{i=1}^n (x_i - C_i(\Omega_e))^2 \right),$$

where the vector $\Phi'$ is defined in (3.16). We will prove there exists a value for the constant $c$ in the expression above such that the matrix $V_{n \times n} + S_{n \times n}$ will be positive.
or negative definite. Finally, using the explicit construction for $\phi$ above, we apply the energy-Casimir method and prove that a generic equilibrium is Lyapunov stable if either the entries of the inertia matrix $J$ satisfy the ordering

$$\lambda_1 > \lambda_3, \lambda_4, \ldots, \lambda_{2n}$$
$$\lambda_2 > \lambda_3, \lambda_4, \ldots, \lambda_{2n}$$
$$\lambda_3 > \lambda_5, \lambda_6, \ldots, \lambda_{2n}$$
$$\lambda_4 > \lambda_5, \lambda_6, \ldots, \lambda_{2n}$$
$$\vdots$$
$$\lambda_{2n-3} > \lambda_{2n-1}, \lambda_{2n}$$
$$\lambda_{2n-2} > \lambda_{2n-1}, \lambda_{2n},$$

or the ordering

$$\lambda_1 < \lambda_3, \lambda_4, \ldots, \lambda_{2n}$$
$$\lambda_2 < \lambda_3, \lambda_4, \ldots, \lambda_{2n}$$
$$\lambda_3 < \lambda_5, \lambda_6, \ldots, \lambda_{2n}$$
$$\lambda_4 < \lambda_5, \lambda_6, \ldots, \lambda_{2n}$$
$$\vdots$$
$$\lambda_{2n-3} < \lambda_{2n-1}, \lambda_{2n}$$
$$\lambda_{2n-2} < \lambda_{2n-1}, \lambda_{2n}.$$

3.1 A General Family of Casimirs

First we find a define family of Casimirs from which we will construct a possible candidate for a Lyapunov function to analyze generic equilibria (2.19) of the rigid body. As noted in Section 2.1, a generic adjoint orbit in $\mathfrak{so}(2n)$ has dimension $\dim(\mathfrak{g}) - \text{rank}(\mathfrak{g}) = 2n^2 - 2n$. A generic adjoint orbit $\sigma_{\Omega_c}$ can be described as a
level set of $n$ linearly independent Casimirs which we previously described in (2.2):

\[ C_1(\Omega) = -\frac{1}{4}\tr (\Omega^2) \]
\[ C_2(\Omega) = \frac{1}{8}\tr (\Omega^4) \]
\[ \vdots \]
\[ C_k(\Omega) = \frac{(-1)^k}{4k}\tr (\Omega^{2k}) \]
\[ \vdots \]
\[ C_{n-1}(\Omega) = \frac{(-1)^{n-1}}{4(n-1)}\tr (\Omega^{2n-2}) \]
\[ C_n = \text{Pf}(\Omega) = \sqrt{\det(\Omega)} \]

where the last Pfaffian $C_n$ is clearly distinct from the other $n-1$ trace Casimirs and will be treated as a special case in much of this chapter.

Eventually it will be useful to evaluate each Casimir at a generic equilibrium (2.19),

\[ C_k(\Omega_e) = \frac{1}{2k} \left(a_1^{2k} + a_2^{2k} + \ldots + a_n^{2k}\right) \quad \text{for } 1 \leq k \leq n-1 \]
\[ C_n(\Omega_e) = a_1 a_2 \ldots a_n. \]  

(3.1)

The factor $(-1)^k$ before the trace Casimirs in (2.2) ensures that all of the sums in (3.1) have positive terms, and the factor $\frac{1}{4k}$ will eventually result in much cleaner derivatives $dC_k(\Omega_e)$ for $1 \leq k \leq n-1$ when computing the first variation of $H_\phi$ at $\Omega_e$. Next we introduce an arbitrary smooth, real-valued function (to be determined later)

\[ \phi : \mathbb{R}^n \rightarrow \mathbb{R} \]  

(3.2)

and the function

\[ F : \mathfrak{so}(2n) \rightarrow \mathbb{R}^n : \Omega \mapsto (C_1(\Omega), C_2(\Omega), \ldots, C_n(\Omega)) \].  

(3.3)

Composing the two functions above will create a general family of Casimirs:

\[ \phi_F : \mathfrak{so}(2n) \rightarrow \mathbb{R} : \Omega \mapsto \phi(C_1(\Omega), C_2(\Omega), \ldots, C_n(\Omega)). \]  

(3.4)
Since the generic adjoint orbits are level sets of the Casimirs in (2.2), the restriction of $F$ to an adjoint orbit $\sigma_{\Omega_e}$ is constant,

$$F|_{\sigma_{\Omega_e}} (\Omega) = (k_1, k_2, \ldots, k_n),$$

for all $\Omega \in \sigma_{\Omega_e}$ and $(k_1, k_2, \ldots, k_n)$ the fixed point in $\mathbb{R}^n$ which defines the level set. Thus it clearly follows that composition of the arbitrary smooth function $\phi$ with $F$ is constant on a fixed generic adjoint orbit $\sigma_{\Omega_e}$. Indeed, $\phi_F$ is a suitable arbitrary Casimir function from which we can define a modified Hamiltonian function and candidate for a Lyapunov function

$$H_{\phi}(\Omega) = H(\Omega) + \phi_F(\Omega). \quad (3.5)$$

### 3.2 The First Variation

Using $H_{\phi}$ in (3.5), we now use Lyapunov’s method to determine sufficient conditions for the stability of generic equilibria $\Omega_e$. The equilibrium $\Omega_e$ must first be a critical point of $H_{\phi}$. We now compute the first variation of $H_{\phi}$ and find conditions on the function $\phi$ that ensure $dH_{\phi}(\Omega_e) = 0$. For any $X$ and $Y \in \mathfrak{so}(2n)$, we now denote the following form (which is derived from the Killing form):\[\langle X, Y \rangle = -\frac{1}{2} \text{tr}(XY). \quad (3.6)\]

**Lemma 3.2.1** Recall the $2n \times 2n$ inertia matrix $J = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{2n})$ (1.36) corresponding to a $2n$ dimensional rigid body. The first variation of the original Hamiltonian $H(\Omega) = -\frac{1}{2} \text{tr}(J\Omega^2)$ is

$$dH(\Omega) = J\Omega + \Omega J \quad (3.7)$$

**Proof** First recall the definition for the first variation $dG(\Omega)$ of a general smooth
function $G : \mathfrak{so}(2n) \rightarrow \mathbb{R}$ at a point $\Omega$ in the Lie algebra $\mathfrak{so}(2n)$

$$
\langle Y, dG(\Omega) \rangle = \left. \frac{d}{dt} \right|_{t=0} G(\Omega + tY) : \quad (3.8)
$$

for any $Y \in \mathfrak{so}(2n)$ and the form $< , >$ defined in (3.6). Using the first variation, we compute the first variation of the Hamiltonian $H_\phi$ defined in (3.5)

$$
\left. \frac{d}{dt} \right|_{t=0} H(\Omega + tY) = \left. \frac{d}{dt} \right|_{t=0} \left[ -\frac{1}{2} \text{tr} \left( J(\Omega + tY)^2 \right) \right]
$$

$$
= \left. \frac{d}{dt} \right|_{t=0} \left\{ -\frac{1}{2} \text{tr} \left[ J \left( \Omega^2 + t(\Omega Y + Y\Omega) + t^2 Y^2 \right) \right] \right\}
$$

$$
= -\frac{1}{2} \text{tr} (J(\Omega Y + Y\Omega))
$$

$$
= -\frac{1}{2} \text{tr} (Y(J\Omega + \Omega J))
$$

$$
= \langle Y, J\Omega + \Omega J \rangle .
$$

In particular, at a generic equilibrium $\Omega_e$

$$
dH(\Omega_e) = J\Omega_e + \Omega_e J
$$

$$
= \left( \begin{array}{cccc}
J_1 H_1 + H_1 J_1 & 0 & \ldots & 0 \\
0 & J_2 H_2 + H_2 J_2 & \vdots & \\
\vdots & \ddots & \ddots & \\
0 & \ldots & 0 & J_n H_n + H_n J_n \\
\end{array} \right) , \quad (3.9)
$$

where the matrix in (3.9) is a $2n \times 2n$ block diagonal matrix with $n$ blocks along the diagonal

$$
J_k H_k + H_k J_k = \left( \begin{array}{cc}
0 & (\lambda_{2k-1} + \lambda_{2k}) a_k \\
-(\lambda_{2k-1} + \lambda_{2k}) a_k & 0 \\
\end{array} \right) \quad \text{for} \quad 1 \leq k \leq n.
$$
Lemma 3.2.2 For \( 1 \leq k \leq n - 1 \), the first variation of \( C_k \) is

\[
dC_k(\Omega) = (-1)^{k-1}\Omega^{2k-1}.
\]

Proof According to the definition of the first variation in (3.8), we compute

\[
\frac{d}{dt} \bigg|_{t=0} C_k(\Omega + tY) = \frac{(-1)^k}{4k} \frac{d}{dt} \bigg|_{t=0} \text{tr}((\Omega + tY)^{2k})
\]

\[
= \frac{(-1)^k}{4k} (2k) \text{tr}(Y\Omega^{2k-1})
\]

\[
= (-1)^{k-1} \left( -\frac{1}{2} \text{tr}(Y\Omega^{2k-1}) \right)
\]

\[
= (-1)^{k-1} \langle Y, \Omega^{2k-1} \rangle.
\]

Finally it follows that \( dC_k(\Omega) = (-1)^{k-1}\Omega^{2k-1} \).

In particular, at a generic equilibrium,

\[
dC_k(\Omega_e) = (-1)^{k-1}\Omega_e^{2k-1}
\]

\[
= \begin{pmatrix}
0 & a_1^{2k-1} & 0 & 0 & \cdots & 0 & 0 \\
-a_1^{2k-1} & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & a_2^{2k-1} & \cdots & 0 & 0 \\
0 & 0 & -a_2^{2k-1} & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & a_{n}^{2k-1} \\
0 & 0 & 0 & 0 & \cdots & -a_{n}^{2k-1} & 0
\end{pmatrix},
\]

(3.11)

and as a result of the leading factor \( \frac{(-1)^k}{4k} \) in (2.2), the entries in (3.11) are particularly nice.
Lemma 3.2.3 The first variation of the Casimir $C_n(\Omega) = \text{Pf}(\Omega)$ is

\[
dC_n(\Omega) = -\left(\sqrt{\det(\Omega)}\right)\Omega^{-1} = -\left(\text{Pf}(\Omega)\right)\Omega^{-1}. \tag{3.12}
\]

Proof First recall, that for $M \in GL_n(\mathbb{R})$,

\[
\left.\frac{d}{dt}\right|_{t=0} (\det(M + Bt)) = \det(M) \text{ tr} (BM^{-1})
\]

for $B \in GL_n(\mathbb{R})$. Using this derivative, again we begin by computing the first variation of $C_n$ at $\Omega$:

\[
\left.\frac{d}{dt}\right|_{t=0} C_n(\Omega + tY) = \left.\frac{1}{2\sqrt{\det(\Omega)}}\right|_{t=0} \left[\frac{d}{dt}\right|_{t=0} \sqrt{\det(\Omega + tY)}
\]
\[
= \frac{1}{2\sqrt{\det(\Omega)}} \cdot \det(\Omega) \cdot \text{ tr} (Y\Omega^{-1})
\]
\[
= -\frac{1}{2} \text{ tr} \left(Y\left(-\sqrt{\det(\Omega)} \cdot \Omega^{-1}\right)\right)
\]
\[
= \left\langle Y, -\sqrt{\det(\Omega)} \cdot \Omega^{-1}\right\rangle,
\]

and indeed $dC_n(\Omega) = -\left(\sqrt{\det(\Omega)}\right)\Omega^{-1}$. ■
Again we evaluate $dC_n(\Omega)$ at $\Omega_e$:

$$dC_n(\Omega_e) = - \left( \sqrt{\det(\Omega_e)} \right) \Omega_e^{-1}$$

$$= - (a_1a_2 \ldots a_n) \begin{pmatrix}
0 & -\frac{1}{a_1} & \ldots & 0 & 0 \\
\frac{1}{a_1} & 0 & \ldots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & -\frac{1}{a_n} & 0 \\
0 & 0 & \ldots & \frac{1}{a_n} & 0
\end{pmatrix}$$

$$= \begin{pmatrix}
\hat{A}_1 & 0 & \ldots & 0 \\
0 & \hat{A}_2 & 0 & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
0 & \ldots & 0 & \hat{A}_n
\end{pmatrix}, \quad \text{(3.14)}$$

where the matrix in (3.14) is a $2n \times 2n$ block diagonal matrix with $n$ blocks along the diagonal,

$$\hat{A}_k = \begin{pmatrix}
0 & a_1a_2 \ldots a_k \ldots a_n \\
-a_1a_2 \ldots a_k \ldots a_n & 0
\end{pmatrix} \quad \text{for } 1 \leq k \leq n.$$
defined in (2.4), and define

$$
\Lambda = \begin{pmatrix}
\lambda_1 + \lambda_2 \\
\lambda_3 + \lambda_4 \\
\vdots \\
\lambda_{2n-1} + \lambda_{2n}
\end{pmatrix},
\Phi' = \begin{pmatrix}
\phi_1(F(\Omega_e)) \\
\vdots \\
\phi_n(F(\Omega_e))
\end{pmatrix},
A = \begin{pmatrix}
a_1 & 0 & \ldots & 0 \\
0 & a_2 & \ldots & 0 \\
& & \ddots & \\
0 & \ldots & 0 & a_n
\end{pmatrix},
$$

(3.15)

d\tilde{F}_{\Omega_e} = Jacobian of \tilde{F} at \Omega_e.

The equilibrium $\Omega_e$ in (2.19) is a critical point of $H_\phi$ if and only if the first partial derivatives of $\phi$ at $F(\Omega_e)$ satisfy the following condition:

$$
\Phi' = -\left( d\tilde{F}_{\Omega_e}' \right)^{-1} A \Lambda
$$

(3.16)

**Proof** The first variation of $H_\phi$ can be split into the sum of two derivatives, $dH_\phi(\Omega_e) = dH(\Omega_e) + d[\phi(F(\Omega_e))]$. The first summand is obtained from Lemma 3.2.1. In order to compute the derivative of $\phi(F(\Omega))$ at $\Omega = \Omega_e$ we need to use the chain rule

$$
d[\phi(F(\Omega_e))] = \phi_1(F(\Omega_e)) dC_1(\Omega_e) + \ldots + \phi_n(F(\Omega_e)) dC_n(\Omega_e).
$$

(3.17)

We wish to solve the equation $dH_\phi(\Omega_e) = 0$. As a result of Lemmas 3.2.2 and 3.2.3, as well as formulas (3.9), (3.11), and (3.14)), the expression (3.17) yields the following system of equations:

$$
\begin{align*}
(\lambda_1 + \lambda_2) a_1 + & \sum_{k=1}^{n-1} [(a_{1k}^{2k-1}) \phi_k(F(\Omega_e))] + (a_2a_3 \ldots a_n) \phi_n(F(\Omega_e)) = 0 \\
(\lambda_3 + \lambda_4) a_2 + & \sum_{k=1}^{n-1} [(a_{2k}^{2k-1}) \phi_k(F(\Omega_e))] + (a_1a_3 \ldots a_n) \phi_n(F(\Omega_e)) = 0 \\
& \vdots \\
(\lambda_{2n-1} + \lambda_{2n}) a_n + & \sum_{k=1}^{n-1} [(a_{nk}^{2k-1}) \phi_k(F(\Omega_e))] + (a_1a_2 \ldots a_{n-1}) \phi_n(F(\Omega_e)) = 0
\end{align*}
$$

(3.18)
Notice that this system can be expressed as a matrix equation:

\[
\begin{pmatrix}
a_1 & a^3_1 & \cdots & a_2 a_3 \ldots a_n \\
a_2 & a^3_2 & \cdots & a_1 a_3 \ldots a_n \\
\vdots & \vdots & & \vdots \\
a_n & a^3_n & \cdots & a_1 a_2 \ldots a_{n-1}
\end{pmatrix}
\begin{pmatrix}
\phi_1 (F(\Omega_e)) \\
\phi_2 (F(\Omega_e)) \\
\vdots \\
\phi_n (F(\Omega_e))
\end{pmatrix}
= 
\begin{pmatrix}
-a_1 & 0 & \cdots & 0 \\
0 & -a_2 & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
0 & 0 & \cdots & -a_n
\end{pmatrix}
\begin{pmatrix}
\lambda_1 + \lambda_2 \\
\lambda_3 + \lambda_4 \\
\vdots \\
\lambda_{2n-1} + \lambda_{2n}
\end{pmatrix}
\] (3.19)

Using the notation introduced in (3.15), equation (3.19) can be written more succinctly as \(d\tilde{F}_t^\epsilon \Phi' = -\Lambda \Lambda\). Since \(\Omega_e\) is a generic equilibrium on a non-degenerate submanifold \(\sigma_{\Omega_e}\), it is a regular point of \(F\) as well as of the restriction \(\tilde{F}\). The Jacobian \(d\tilde{F}\) is therefore invertible at \(\Omega_e\) (as well as the transpose of the Jacobian). Finally, by applying the inverse of \(d\tilde{F}_t^\epsilon\) to both sides, we have proved (3.16).

### 3.3 The Second Variation

Continuing with the energy-Casimir method, we will now compute the second variation at \(\Omega_e\),

\[
d^2 H_\phi (\Omega_e) (Z, Y) = \left\langle Z, \left. \frac{d}{dt} \right|_{t=0} dH_\phi (\Omega_e + tY) \right\rangle,
\] (3.20)

for any pair of matrices \(Y, Z \in so(2n)\). If a function \(\phi\) can be defined such that (3.16) is satisfied and \(d^2 H_\phi (\Omega_e)\) is positive or negative definite, then a generic equilibrium \(\Omega_e\) is Lyapunov stable. We will proceed with this step by first showing the second variation can be represented by a block diagonal matrix consisting of one large \(n \times n\) block and a total of \(n^2 - n\) blocks each of size \(2 \times 2\). Altogether the matrix \(d^2 H_\phi (\Omega_e)\) will therefore be a \((2n^2 - n) \times (2n^2 - n)\) symmetric, block diagonal matrix. Definiteness of the second variation will therefore be equivalent to definiteness of the block diagonal matrix.

**Proposition 3.3.1** The second variation of \(H_\phi\) at \(\Omega_e\), applied to a pair of matrices \(Y, Z \in so(2n)\), can be expressed by

\[
d^2 H_\phi (\Omega_e)(Z, Y) = d^2 H(\Omega_e)(Z, Y) + First(\Omega_e) + Second(\Omega_e),
\] (3.21)
where First(\(\Omega_e\)) and Second(\(\Omega_e\)) denote sums of terms involving first partial derivatives \(\phi_l(F(\Omega_e))\) (for \(1 \leq l \leq n\)) and second partial derivatives \(\phi_{lk}(F(\Omega_e))\) (for \(1 \leq l, k \leq n\)) respectively:

\[
\begin{align*}
\text{First}(\Omega_e) &= \sum_{l=1}^{n} \{ \phi_l(F(\Omega_e)) \, d^2C_l(\Omega_e)(Z,Y) \} \\
\text{Second}(\Omega_e) &= \sum_{l=1}^{n} \sum_{k=1}^{n} \{ \phi_{lk}(F(\Omega_e)) \, \langle Y, dC_k(\Omega_e) \rangle \langle Z, dC_l(\Omega_e) \rangle \} \\
\end{align*}
\]

(3.22)

Proof We begin by computing \(d^2H_{\phi}(\Omega_e)\) using the definition of the second variation (3.20):

\[
\begin{align*}
d^2H_{\phi}(\Omega_e)(Z,Y) &= \langle Z, \frac{d}{dt}|_{t=0} dH_{\phi}(\Omega_e + tY) \rangle \\
&= d^2H(\Omega_e)(Z,Y) + \langle Z, \frac{d}{dt}|_{t=0} d[\phi(F(\Omega_e + tY))] \rangle.
\end{align*}
\]

(3.23)

Using the standard chain and product rules along with the expression (3.17) applied to \(d[\phi(F(\Omega_e + tY))]\), we differentiate the second term in (3.23):

\[
\begin{align*}
\frac{d}{dt}|_{t=0} (d[\phi(F(\Omega_e + tY))]) &= \frac{d}{dt}|_{t=0} \{ \sum_{l=1}^{n} \{ \phi_l(F(\Omega_e + tY)) \, dC_l(\Omega_e + tY) \} \} \\
&= \sum_{l=1}^{n} \sum_{k=1}^{n} \{ \phi_{lk}(F(\Omega_e)) \, \langle Y, dC_k(\Omega_e) \rangle \cdot dC_l(\Omega_e) \} \\
&+ \sum_{l=1}^{n} \{ \phi_l(F(\Omega_e)) \cdot \frac{d}{dt}|_{t=0} dC_l(\Omega_e + tY) \}.
\end{align*}
\]

(3.24)

Substituting (3.24) back into (3.23),

\[
\begin{align*}
\langle Z, \frac{d}{dt}|_{t=0} d[\phi(F(\Omega_e + tY))] \rangle &= \sum_{l=1}^{n} \sum_{k=1}^{n} \{ \phi_{lk}(F(\Omega_e)) \, \langle Y, dC_k(\Omega_e) \rangle \, \langle Z, dC_l(\Omega_e) \rangle \} \\
&+ \sum_{l=1}^{n} \{ \phi_l(F(\Omega_e)) \, \langle Z, \frac{d}{dt}|_{t=0} dC_l(\Omega_e + tY) \rangle \} \\
&= \text{First}(\Omega_e) + \text{Second}(\Omega_e).
\end{align*}
\]

3.3.1 A Basis for Setting Up a Matrix Representation for \(d^2H_{\phi}(\Omega_e)\)

If Proposition 3.3.1 tells us anything, it is that we should find a friendlier expression for the second variation. We will construct a symmetric matrix to represent the
bilinear form in (3.21). This takes three steps: first we will compute a matrix to represent $d^2H(\Omega_e)$, next we will construct a matrix to represent \textit{First}(\Omega_e) in (3.22), and lastly we will construct a matrix for \textit{Second}(\Omega_e) in (3.22). Clearly, the structure of the matrix depends on the basis we choose to represent $so(2n)$. In this section we will define a basis in which the matrix will be block diagonal. Such a form will be especially helpful in establishing conditions for definiteness of the second variation.

Before we attempt to define a basis for $so(2n)$ in general, we consider first the simplest case when $g = so(4)$. Let $\Omega = (\omega_{ij}) \in so(4)$. We introduce coordinates $\tilde{y}_j$ as follows:

$$
\Omega = \begin{pmatrix}
  0 & \omega_{12} & \omega_{13} & \omega_{14} \\
-\omega_{12} & 0 & \omega_{23} & \omega_{24} \\
-\omega_{13} & -\omega_{23} & 0 & \omega_{34} \\
-\omega_{14} & -\omega_{24} & -\omega_{34} & 0 \\
\end{pmatrix} = \begin{pmatrix}
  0 & \tilde{y}_1 & \tilde{y}_3 & \tilde{y}_5 \\
-\tilde{y}_1 & 0 & \tilde{y}_6 & \tilde{y}_4 \\
-\tilde{y}_3 & -\tilde{y}_6 & 0 & \tilde{y}_2 \\
-\tilde{y}_5 & -\tilde{y}_4 & -\tilde{y}_2 & 0 \\
\end{pmatrix}
$$

(3.25)

Notice that in (3.25) the $\tilde{y}_j$ are indexed in an unconventional manner. This order is chosen so the matrix representing $d^2H_\phi(\Omega_e)$ will in fact be block diagonal. At the same time the natural row/column indices of the entries $\omega_{ij}$ will be much more useful when setting up the matrix. Similarly consider the next case, $so(6)$, to further illustrate the ordering:

$$
\Omega = \begin{pmatrix}
  0 & \tilde{y}_1 & \tilde{y}_4 & \tilde{y}_6 & \tilde{y}_8 & \tilde{y}_{10} \\
-\tilde{y}_1 & 0 & \tilde{y}_7 & \tilde{y}_5 & \tilde{y}_{11} & \tilde{y}_9 \\
-\tilde{y}_4 & -\tilde{y}_7 & 0 & \tilde{y}_2 & \tilde{y}_{12} & \tilde{y}_{14} \\
-\tilde{y}_6 & -\tilde{y}_5 & -\tilde{y}_2 & 0 & \tilde{y}_{15} & \tilde{y}_{13} \\
-\tilde{y}_8 & -\tilde{y}_{11} & -\tilde{y}_{12} & -\tilde{y}_{15} & 0 & \tilde{y}_3 \\
-\tilde{y}_{10} & -\tilde{y}_9 & -\tilde{y}_{14} & -\tilde{y}_{13} & -\tilde{y}_3 & 0 \\
\end{pmatrix}
$$

In general we continue to label entries of $\Omega \in so(2n)$ in the same fashion. We begin indexing entries along the diagonal blocks of $\Omega$ and proceed to the off-diagonal
blocks. In each off-diagonal block, we label diagonal entries of the block first, followed by the anti-diagonal entries of the block. Now let \( \tilde{Y}_i \) denote the matrix with entries \( \tilde{y}_i = 1 \) and \( \tilde{y}_j = 0 \) for \( j \neq i \). The collection of matrices \( \{ \tilde{Y}_i \} \) with \( 1 \leq i \leq (2n^2 - n) = \dim(\mathfrak{so}(2n)) \) defines a basis for \( \mathfrak{so}(2n) \). With respect to this basis, \( d^2H_\phi(\Omega_e) \) becomes a symmetric (and later block diagonal) matrix with \( i_j \)th entry

\[
(d^2H_\phi(\Omega_e))_{ij} = d^2H_\phi(\Omega_e) \left( \tilde{Y}_i, \tilde{Y}_j \right) = \left. \frac{d}{dt} \right|_{t=0} \left( dH_\phi(\Omega_e + t\tilde{Y}_j) \right).
\]

(3.26)

Remark 3.3.2 For future reference we evaluate the trace form \( \langle \cdot, \cdot \rangle \) in (3.6) with respect to the basis \( \{ \tilde{Y}_i \} \) at a generic equilibrium (2.19):

\[
\langle \tilde{Y}_i, \tilde{Y}_j \rangle = \delta_{ij}
\]

\[
\langle \tilde{Y}_i, \Omega_e \rangle = \begin{cases} a_i & \text{if } \tilde{Y}_i \text{ is a diagonal basis element} \\ 0 & \text{otherwise} \end{cases}
\]

(3.27)

As previously noted, it will be more practical to know exactly which entries of each \( \tilde{Y}_i \) (in the usual row/column convention) are non-zero. A general matrix \( \Omega \in \mathfrak{so}(2n) \) can be divided into \( n^2 \) distinct \( 2 \times 2 \) blocks:

\[
\Omega = \begin{pmatrix}
  h_1 & M_{12} & M_{13} & \ldots & M_{1n} \\
  h_2 & M_{23} & \ldots & M_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  h_{n-1} & M_{n-1,2} & \ldots & h_n
\end{pmatrix}
\]

(3.28)

Note that there are \( \binom{n}{2} = \frac{n(n-1)}{2} \) blocks \( M_{r,c} \) in the upper triangle. The indices \( r \)
and $c$ indicate that the block $M_{r,c}$ is in the $r^{\text{th}}$ row and $c^{\text{th}}$ column of $2 \times 2$ matrices. The $h_j$'s span the Cartan subalgebra $\mathfrak{h}$ defined in (2.5) of Section 2.4. We use the same notation as in (2.4) to denote the diagonal basis elements of $\{Y_i\}$:

$$Y^j_d = e_{jj} \otimes i\sigma_2 = \tilde{Y}_j \quad \text{for} \quad 1 \leq j \leq n. \quad (3.29)$$

The remaining off-diagonal basis elements are labelled quite differently. Each $2 \times 2$ block $M_{r,c}$ is spanned by four distinct basis elements corresponding to the four entries in the each block. We now define the four basis elements $Y^{e_{ij}}_{r,c}$ which span each block $M_{r,c}$ as follows:

$$Y^{e_{ij}}_{r,c} = \begin{cases} 
\text{all} \ 2 \times 2 \ \text{blocks of} \ \Omega \ \text{(as divided in (3.28)) are trivial} \\
\text{except the block} \ M_{r,c} = e_{ij} \ \text{for} \ 1 \leq i, j \leq 2, 
\end{cases} \quad (3.30)$$

where $e_{ij}$ denotes the canonical basis element for $\text{Mat}_{2 \times 2}(\mathbb{R})$. For example, if $g = \mathfrak{so}(6)$, we would write

$$Y^{e_{21}}_{1,3} = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 
\end{pmatrix}. \quad (3.31)$$

Lastly, we describe the matrices $Y^i_d$ and $Y^{e_{ij}}_{r,c}$ in the parameterization of $\Omega$ by $\omega_{ij}$. Each diagonal matrix $Y^i_d$ corresponds to the matrix with all zero entries except $\omega_{(2i-1,2i)} = 1 = -\omega_{(2i,2i-1)}$. Next, the block $M_{r,c}$ has the form

$$M_{r,c} = \begin{pmatrix}
\omega(2r-1,2c-1) & \omega(2r-1,2c) \\
\omega(2r,2c-1) & \omega(2r,2c) 
\end{pmatrix}. \quad (3.31)$$
Therefore the off-diagonal basis elements $Y_{r,c}^{e_{ij}}$ correspond to matrices with all entries zero except possibly in $M_{r,c}$:

$$Y_{r,c}^{e_{ij}} = \begin{cases} 
\omega(2r-1,2c-1) = 1 = -\omega(2c-1,2r-1) & e_{ij} = e_{11} \\
\omega(2r-1,2c) = 1 = -\omega(2c,2r-1) & e_{ij} = e_{12} \\
\omega(2r-1,2c-1) = 1 = -\omega(2c-1,2r) & e_{ij} = e_{21} \\
\omega(2r,2c) = 1 = -\omega(2c,2r) & e_{ij} = e_{22}
\end{cases} \quad (3.32)$$

We could similarly find a relationship between the $\omega_{ij}$’s and the corresponding indices of the basis elements $Y_i$; however, this relationship is quite complicated and not particularly helpful.

Before moving on to the construction of the matrix $d^2H_\phi(\Omega_e)$, we make several observations regarding the notation introduced in this section which will play a fundamental role. Refer to equation (3.68) of Corollary 3.3.9 at the end of this section for the final block diagonal expression of $d^2H_\phi(\Omega_e)$.

- Evaluating the second variation at a pair of diagonal basis elements $d^2H_\phi(\Omega_e)(Y_{d,i}^i, Y_{d,j}^j)$ (for $1 \leq i, j \leq n$) will give rise to entries in the first $n \times n$ block of $d^2H_\phi(\Omega_e)$ we will eventually denote $V_{n \times n} + S_{n \times n}$ (3.68).
- Evaluating the second variation at a pair $d^2H_\phi(\Omega_e)(Y_{r,c}^{e_{ij}}, Y_{r,c}^{e_{ji}})$ for all $1 \leq i, j \leq 2$ will give rise to the $2 \times 2$ blocks of $d^2H_\phi(\Omega_e)$ we will call $V_{2 \times 2}^{diag}$ (3.68).
- Likewise evaluating $d^2H_\phi(\Omega_e)(Y_{r,c}^{e_{ij}}, Y_{r,c}^{e_{ji}})$, for all $1 \leq i, j \leq 2$ and $i \neq j$, will give rise to the $2 \times 2$ blocks we will denote $V_{2 \times 2}^{off}$ (3.68).
- The ordering of the pairs of consecutive blocks $V_{r,c}^{diag}$ and $V_{r,c}^{off}$ (for $1 \leq r < c \leq n$) along the diagonal of $d^2H_\phi(\Omega_e)$ is a result of the order in which we chose to index the off-diagonal blocks $M_{r,c}$. For example we could have chosen to index the block $M_{n-1,n}$ first, followed by $M_{1,2}$, etcetera. In which case, the
pair \( V_{n-1,n}^{\text{diag}} \) and \( V_{n-1,n}^{\text{off}} \) would appear first in (3.68), followed by the pair \( V_{1,2}^{\text{diag}} \) and \( V_{1,2}^{\text{off}} \), and so on.

### 3.3.2 Constructing the Matrix of \( d^2 H \)

We now begin the process of evaluating the entries of \( d^2 H_\phi(\Omega_c) \) with respect to the basis \( \{\tilde{Y}_i\} \). Returning to Proposition 3.3.1 we proceed first with \( d^2 H(\Omega_c) \).

**Theorem 3.3.3** In the basis \( \{\tilde{Y}_i\} \), the second variation of \( H \) at \( \Omega_c \) is represented by the \((2n^2 - n) \times (2n^2 - n)\) diagonal matrix

\[
d^2 H(\Omega_c) = \begin{pmatrix}
\lambda_1 + \lambda_2 & 0 & \ldots & 0 \\
0 & \lambda_3 + \lambda_4 & & \\
& \ddots & \ddots & \\
0 & \ldots & 0 & \lambda_{2n-1} + \lambda_{2n}
\end{pmatrix}
\]

\[
L_{1,2}^{\text{diag}} \\
L_{1,2}^{\text{off}} \\
\ddots \\
L_{n-1,1}^{\text{off}}
\]

(3.33)

where in the \( 2 \times 2 \) diagonal matrices \( L_{r,c}^{\text{diag}} \) and \( L_{r,c}^{\text{off}} \) have the form

\[
L_{r,c}^{\text{diag}} = \begin{pmatrix}
\lambda_{2r-1} + \lambda_{2c-1} & 0 \\
0 & \lambda_{2r} + \lambda_{2c}
\end{pmatrix}
\]

\[
L_{r,c}^{\text{off}} = \begin{pmatrix}
\lambda_{2r-1} + \lambda_{2c} & 0 \\
0 & \lambda_{2r} + \lambda_{2c-1}
\end{pmatrix}
\]

(3.34)

for all \( 1 \leq r < c \leq n \).

**Proof** Recall from Lemma 3.2.1 that \( dH(\Omega_c + t\tilde{Y}_j) = J \left( \Omega_c + t\tilde{Y}_j \right) + \left( \Omega_c + t\tilde{Y}_j \right) J \).
Thus,

\[
d^2 H(\Omega_e)(\tilde{Y}_i, \tilde{Y}_j) = \left. \left. \frac{d}{dt} \right|_{t=0} \left( \frac{d}{dt} \left( \Omega_e + t\tilde{Y}_j \right) \right) \right| \tilde{Y}_i, \tilde{Y}_j \right.
\]

\[
= \left. \frac{d}{dt} \left( J(\Omega_e + t\tilde{Y}_j) + (\Omega_e + t\tilde{Y}_j)J \right) \right|_{t=0}
\]

\[
= \left( \tilde{Y}_i, J\tilde{Y}_j + \tilde{Y}_j J \right)
\]

\[
= -\frac{1}{2} \text{tr} \left( \tilde{Y}_i (J\tilde{Y}_j + \tilde{Y}_j J) \right)
\]

\[
= -\frac{1}{2} \text{tr} \left( J(\tilde{Y}_j \tilde{Y}_i + \tilde{Y}_i \tilde{Y}_j) \right). \tag{3.35}
\]

First observe if \( \tilde{Y}_i \neq \tilde{Y}_j \) the product \( \tilde{Y}_j \tilde{Y}_i = \tilde{Y}_i \tilde{Y}_j = 0 \), and thus (3.35) vanishes. Indeed, the matrix \( d^2 H(\Omega_e) \) is diagonal. Now we focus our attention solely on pairs of equal matrices \( \tilde{Y}_i = \tilde{Y}_j \) which correspond to the diagonal entries of \( d^2 H(\Omega_e) \).

Let \( \tilde{Y}_i = \tilde{Y}_j = Y_d^i \) be diagonal basis elements and compute \( d^2 H(\Omega_e)(Y_d^i, Y_d^i) \)

\[
-\frac{1}{2} \text{tr} \left( Y_d^i (JY_d^i + Y_d^i J) \right) = -\text{tr} \left( (JY_d^i)^2 \right). 
\]

The square of the matrix \( Y_d^i \) is a \( 2n \times 2n \) diagonal matrix with entries \( \omega_{(2i-1,2i-1)} = \omega_{(2i,2i)} = -1 \) and 0 otherwise. The product of the diagonal matrices \( J \) and \( (Y_d^i)^2 \) is a diagonal matrix \( D \) with \( d_{2i-1} = -\lambda_{2i-1}, d_{2i} = -\lambda_{2i}, \) and \( d_k = 0 \) otherwise; therefore, substituting \( 2D = 2J(Y_d^i)^2 \) into (3.35) we obtain

\[
d^2 H(\Omega_e)(Y_d^i, Y_d^i) = -\text{tr} \left( (JY_d^i)^2 \right) = \lambda_{2i-1} + \lambda_{2i}. \tag{3.36}
\]

This yields the first \( n \) diagonal entries of (3.33).

Next we take into account the off-diagonal basis elements \( Y_{r,c}^{e,ij} \). In general, if \( \Omega_{ij} \) is a skew-symmetric matrix with one non-zero entry \( \omega_{ij} = 1 \) in the upper triangle (and thus \( \omega_{ji} = -1 \) below the diagonal), then \( \Omega_{ij}^2 \) is a diagonal matrix with exactly two non-zero diagonal entries: \( \omega_{ii} = \omega_{jj} = -1 \). From this observation, along with
equation (3.32), we can conclude that $(Y_{r,c}^{e_{ij}})^2$ has all zero entries except for the following diagonal entries for each case of $e_{ij}$ below:

$$
(Y_{r,c}^{e_{ij}})^2 = \begin{cases} 
\omega(2r-1,2r-1) = \omega(2c-1,2c-1) = -1 & e_{ij} = e_{11} \\
\omega(2r,2r) = \omega(2c,2c) = -1 & e_{ij} = e_{22} \\
\omega(2r-1,2r-1) = \omega(2c,2c) = -1 & e_{ij} = e_{12} \\
\omega(2r,2r) = \omega(2c-1,2c-1) = -1 & e_{ij} = e_{21} 
\end{cases} \quad (3.37)
$$

Finally, plugging the matrices $(Y_{r,c}^{e_{ij}})^2$ in (3.37) into equation (3.35), we simplify each of the four cases:

$$
d^2 H(\Omega_e) (Y_{r,c}^{e_{ij}}, Y_{r,c}^{e_{ij}}) = \begin{cases} 
\lambda_{2r-1} + \lambda_{2c-1} & e_{ij} = e_{11} \\
\lambda_{2r} + \lambda_{2c} & e_{ij} = e_{22} \\
\lambda_{2r-1} + \lambda_{2c} & e_{ij} = e_{12} \\
\lambda_{2r} + \lambda_{2c-1} & e_{ij} = e_{21} 
\end{cases} \quad (3.38)
$$

Observe in (3.38) for all $(r,c)$ with $1 \leq r \leq c \leq n$, the first two cases give rise to the matrices of the form $L_{r,c}^{\text{diag}}$ in (3.34), while the last two cases give rise to the matrices $L_{r,c}^{\text{off}}$ in (3.34).

3.3.3 The Matrix of First Partial Derivatives

Before stating the next theorem we must introduce some notation. Let $p$ be a positive, even integer and consider the even and odd sums

$$
E_p(a_i, a_j) = a_i^p + a_i^{p-2} a_j^2 + a_i^{p-4} a_j^4 + \ldots + a_j^p \\
O_p(a_i, a_j) = a_i^{p-1} a_j + a_i^{p-3} a_j^3 + \ldots + a_i a_j^{p-1}. \quad (3.39)
$$
Next introduce two $2 \times 2$ symmetric matrices $B_{r,c}^+(k)$ and $B_{r,c}^-(k)$:

$$B_{r,c}^+(k) = \begin{pmatrix} E_{2k-2}(a_r, a_c) & O_{2k-2}(a_r, a_c) \\ O_{2k-2}(a_r, a_c) & E_{2k-2}(a_r, a_c) \end{pmatrix} \quad \text{and}$$

$$B_{r,c}^-(k) = \begin{pmatrix} E_{2k-2}(a_r, a_c) & -O_{2k-2}(a_r, a_c) \\ -O_{2k-2}(a_r, a_c) & E_{2k-2}(a_r, a_c) \end{pmatrix}$$

(3.40)

**Theorem 3.3.4** The second variation of the Casimir $C_k(\Omega)$ (see (2.2)) for $1 \leq k < n-1$ is represented by the following $(2n^2-n) \times (2n^2-n)$ symmetric, block diagonal matrix:

$${d^2C_k(\Omega_e)} = \begin{pmatrix} (2k-1)a_1^{2k-2} & \cdots & 0 & \cdots & \cdots & \cdots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & (2k-1)a_n^{2k-2} & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & a_1^{2k-2} \end{pmatrix} \begin{pmatrix} B_{1,2}^+(k) & \cdots & \cdots \\ \cdots & \ddots & \cdots \\ \cdots & \cdots & B_{1,2}^-(k) \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} B_{2n-1,2n}^+(k) \\ \cdots \\ \cdots \\ \cdots \end{pmatrix}$$

**Remark 3.3.5** Before proving the theorem, we note that there are $n$ diagonal entries $a_i$ in the equilibrium $\Omega_e$ and $\frac{n(n-1)}{2}$ ways of picking 2 entries from the $n$ diagonal entries. For each pair $(a_r, a_c)$, there are two $2 \times 2$ symmetric matrices $B_{r,c}^+(k)$ and $B_{r,c}^-(k)$ along the diagonal of $d^2C_k(\Omega_e)$ above. Thus in total there are $(n^2-n)$ $2 \times 2$ matrices and one $n \times n$ diagonal matrix which in fact yields a $(2n^2-n) \times (2n^2-n)$ symmetric, block diagonal matrix $d^2C_k(\Omega_e)$.

**Proof** For $k = 1$, the second variation $d^2C_1(\Omega_e)$ corresponding to the Casimir $C_1(\Omega) = -\frac{1}{2} \text{tr} (\Omega^2)$ is a somewhat special case since $B_{r,c}^+(1) = B_{r,c}^-(1) = I_2$ and $(2(1)-1)a_1^{2(1)-2} = 1$. Thus for $k = 1$, we must show $d^2C_1(\Omega_e) = I_{2n^2-n}$. Recall from the formula for $dC_k(\Omega)$ given in equation (3.10) that $dC_1(\Omega) = \Omega$ and compute the
The trace relations with respect to the basis \{\tilde{Y}_j\} found in equation (3.27) imply \(d^2C_1(\Omega_e)(\tilde{Y}_i, \tilde{Y}_j) = \delta_{ij}\) and indeed \(d^2C_1(\Omega_e) = I_{2n^2-n}\).

Next we continue to the general case for \(2 \leq k \leq n-1\). We first make several important observations. For a generic equilibrium \(\Omega_e\) in (2.19), we have

\[
(\Omega_e)^p = \begin{pmatrix}
0 & (-1)^{\frac{p+1}{2}} a_1^p & 0 & \cdots & 0 & 0 \\
(-1)^{\frac{p+1}{2}} a_1^p & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & (-1)^{\frac{p+1}{2}} a_2^p & 0 & \cdots & 0 \\
0 & 0 & 0 & (-1)^{\frac{p+1}{2}} a_2^p & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & (-1)^{\frac{p+1}{2}} a_n^p \\
0 & 0 & 0 & 0 & \cdots & 0 & (-1)^{\frac{p+1}{2}} a_n^p \\
\end{pmatrix}
\]

Next set \(h_r = \begin{pmatrix} 0 & a_r \\ -a_r & 0 \end{pmatrix}\) and \(h_c = \begin{pmatrix} 0 & a_c \\ -a_c & 0 \end{pmatrix}\). and let \(e_{ij}\) denote the standard basis for \(Mat_{2 \times 2}(\mathbb{R})\). If \(p\) and \(q\) are two positive, even integers, then \((h_r)^p\) and \((h_c)^q\) are both diagonal. We therefore compute

\[
(h_r)^pe_{ij}(h_c)^q = \begin{pmatrix} (-1)^{\frac{p+1}{2}} a_r^p a_c^q \end{pmatrix} e_{ij}
\]
Similarly if $p$ and $q$ are odd, positive integers, then we note

\[
(h_r)^p e_{11}(h_c)^q = \left( -1 \right)^{\frac{p+q}{2}} a_r^p a_c^q \]
\[
(h_r)^p e_{22}(h_c)^q = \left( -1 \right)^{\frac{p+q}{2}} a_r^p a_c^q \]
\[
(h_r)^p e_{22}(h_c)^q = \left( -1 \right)^{\frac{p+q}{2}} a_r^p a_c^q \]

\[
(3.43)
\]

Returning to construction of the matrix $d^2 C_k(\Omega_e)$, we use the formula for $dC_k(\Omega)$ in (3.10) to compute

\[
\left( d^2 C_k(\Omega_e) \right)_{ij} = d^2 C_k(\Omega_e)(\bar{Y}_i, \bar{Y}_j)
\]
\[
= \left. \frac{d}{dt} \right|_{t=0} dC_k(\Omega_e + t\bar{Y}_j)
\]
\[
= (-1)^{k-1} \left. \frac{d}{dt} \right|_{t=0} (\Omega_e + t\bar{Y}_j)^{2k-1} \]  . (3.44)

Computing the derivative in (3.44) we thus obtain the linear terms in the expansion of $(\Omega_e + t\bar{Y}_j)^{2k-1}$,

\[
\left. \frac{d}{dt} \right|_{t=0} (\Omega_e + t\bar{Y}_j)^{2k-1} = (\Omega_e)^{2k-2}\bar{Y}_j + (\Omega_e)^{2k-3}\bar{Y}_j^2 \Omega_e + \ldots + \bar{Y}_j(\Omega_e)^{2k-2}
\]
\[
= \sum_{l=0}^{k-1} \left[ \Omega_e^{(2k-2l-2)} \bar{Y}_j \Omega_e^{2l} \right] + \sum_{l=0}^{k-2} \left[ \Omega_e^{(2k-2l-3)} \bar{Y}_j^2 \Omega_e^{2l+1} \right] . (3.45)
\]

We need to calculate $d^2 C_k(\Omega_e)$ at all such pairs $(\bar{Y}_i, \bar{Y}_j)$. First set $\bar{Y}_j = Y_{r,c}^{e_11}$, the off-diagonal basis element defined in (3.30). Substituting $Y_{r,c}^{e_11}$ into equation (3.45), we obtain a sum of terms of the type
\[(\Omega_e)^p Y_{r,c}^{e_{11}} (\Omega_e)^q = \begin{pmatrix}
0 & \ldots & 0 & \text{col } c & 0 \\
\vdots & & \tilde{M}_{r,c} & \ldots & 0 \\
\text{row } r & 0 & \ldots & 0 & \text{row } c
\end{pmatrix}
\]

where \(\tilde{M}_{r,c} = (h_r)^p e_{11} (h_c)^q\), \(h_r\) and \(h_c\) the \(r^{th}\) and \(c^{th}\) diagonal blocks of \(\Omega_e\), and \(p + q = 2k - 2\). When \(p\) and \(q\) both even, positive integers such that \(p + q = 2k - 2\) (and thus \(\frac{p+q}{2} = k - 1\)), equation (3.42) gives

\[
\tilde{M}_{r,c} = (h_r)^p e_{11} (h_c)^q = -(1)^{k-1} a_r^p a_c^q e_{11}. \tag{3.47}
\]

On the other hand, if \(p\) and \(q\) are both odd integers which sum to \(2k - 2\), we use (3.43) gives

\[
\tilde{M}_{r,c} = (h_r)^p e_{11} (h_c)^q = (-1)^{k-1} a_r^p a_c^q e_{22}. \tag{3.48}
\]

Using (3.47) and (3.48), when \(\tilde{Y}_j = Y_{r,c}^{e_{11}}\) equation (3.45) becomes

\[
\frac{d}{dt} \bigg|_{t=0} (\Omega_e + t\tilde{Y}_j)^{2k-1} = (-1)^{k-1} \left( E_{2k-2}(a_r, a_c) Y_{r,c}^{e_{11}} + O_{2k-2}(a_r, a_c) Y_{r,c}^{e_{22}} \right),
\]

which can then be substituted into

\[
(d^2 C_k(\Omega_e))_{ij} = d^2 C_k(\Omega_e)(\tilde{Y}_i, \tilde{Y}_j)
\]

\[
= (-1)^{2(k-1)} \left\langle \tilde{Y}_i, E_{2k-2}(a_r, a_c) Y_{r,c}^{e_{11}} + O_{2k-2}(a_r, a_c) Y_{r,c}^{e_{22}} \right\rangle
\]

\[
E_{2k-2}(a_r, a_c) \quad \tilde{Y}_i = \tilde{Y}_j = Y_{r,c}^{e_{11}}
\]

\[
O_{2k-2}(a_r, a_c) \quad \tilde{Y}_i = \tilde{Y}_{j+1} = Y_{r,c}^{e_{22}}.
\]

\[ \text{otherwise} \]
Setting $\tilde{Y}_j = Y^{e11}_{r,c}$ in $d^2C_k \left( \bar{Y}_i, \tilde{Y}_j \right)$ we obtain the $j^{th}$ column of the corresponding matrix $d^2C_k(\Omega_e)$ in (3.3.4). Therefore the calculation in (3.49) implies the $j^{th}$ column has all zero entries, except for the entries $E_{2k-2}(a_r, a_c)$ and $O_{2k-2}(a_r, a_c)$ in rows $j$ and $j+1$ respectively. By symmetry of the form, $d^2C_k \left( \bar{Y}_i, \tilde{Y}_j \right) = d^2C_k \left( \tilde{Y}_j, \bar{Y}_i \right)$, we similarly obtain column $j+1$ of $d^2C_k(\Omega_e)$ has all zero entries except for $O_{2k-2}(a_r, a_c)$ and $E_{2k-2}(a_r, a_c)$ in rows $j$ and $j+1$ respectively. Therefore, for each pair $(r, c)$ we obtain a $2 \times 2$ matrix of the form

$$B^+_{r,c}(k) = \begin{pmatrix} E_{2k-2}(a_r, a_c) & O_{2k-2}(a_r, a_c) \\ O_{2k-2}(a_r, a_c) & E_{2k-2}(a_r, a_c) \end{pmatrix} \quad \text{for } 1 \leq r < c \leq n$$

along the diagonal of $d^2C_k(\Omega_e)$.

The blocks $B^-_{r,c}(k)$ are similarly obtained. The terms $O_{2k-2}(a_r, a_c)$ in this case (when $e_{ij} = e_{12}$) are negative since when $p$ and $q$ are odd, each product $h^p e_{ij} h^q$ has a leading factor of the form $(-1)^{\frac{p+q-2}{2}}$ instead of $(-1)^{\frac{p+q}{2}}$ (see (3.43)). Therefore again setting $\tilde{Y}_j = Y^{e11}_{r,c}$, it follows $\tilde{Y}_{j+2} = Y^{e12}_{r,c}$, and column $j+2$ is determined by the calculation of $d^2C_k \left( \bar{Y}_i, \tilde{Y}_{j+2} \right)$:

$$\left( d^2C_k(\Omega_e) \right)_{i,j+2} = d^2C_k(\Omega_e)(\bar{Y}_i, \tilde{Y}_{j+2}) = (-1)^{2(k-1)} \left\{ \begin{array}{ll} E_{2k-2}(a_r, a_c) & \tilde{Y}_i = \tilde{Y}_{j+2} = Y^{e12}_{r,c} \\ -O_{2k-2}(a_r, a_c) & \tilde{Y}_i = \tilde{Y}_{j+3} = Y^{e21}_{r,c} \\ 0 & \text{otherwise} \end{array} \right. \quad (3.50)$$

And thus column $j+2$ of $d^2C_k(\Omega_e)$ has all zero entries, except $E_{2k-2}(a_r, a_c)$ and $-O_{2k-2}(a_r, a_c)$ in rows $j+2$ and $j+3$ respectively. Again using the symmetry of the form, $d^2C_k \left( \bar{Y}_i, \tilde{Y}_{j+2} \right) = d^2C_k \left( \tilde{Y}_{j+2}, \bar{Y}_i \right)$, we similarly obtain column $j+3$ of $d^2C_k(\Omega_e)$, which has all zero entries except for $-O_{2k-2}(a_r, a_c)$ and $E_{2k-2}(a_r, a_c)$ in
rows $j + 2$ and $j + 3$ respectively. Therefore, for each pair $(r, c)$ we obtain a $2 \times 2$ matrix of the form

$$B_{r,c}^-(k) = \begin{pmatrix} E_{2k-2}(a_r, a_c) & -O_{2k-2}(a_r, a_c) \\ -O_{2k-2}(a_r, a_c) & E_{2k-2}(a_r, a_c) \end{pmatrix} \quad \text{for } 1 \leq r < c \leq n$$

along the diagonal of $d^2C_k(\Omega_e)$. Moreover, the block $B_{r,c}^-(k)$ immediately follows the block $B_{r,c}^+(k)$ we derived using the calculation in (3.49).

Now that we have shown that there are indeed $(n^2 - n) / 2 \times 2$ blocks along the diagonal of the form $B_{r,c}^-(k)$ and $B_{r,c}^+(k)$, we next verify that first $n \times n$ diagonal matrix has entries $(2k - 1)a_i^{2k-2}$ for $1 \leq i \leq n$. This block arises from evaluating $d^2C_k(\Omega_e)$ at pairs $(\tilde{Y}_i, \tilde{Y}_j)$ for $\tilde{Y}_i = Y_i = Y_i^d$, the diagonal basis matrix defined in (3.29). In order to compute $d^2C_k(\Omega_e)(\tilde{Y}_i, \tilde{Y}_j)$ let $\tilde{Y}_j = Y_j^d$, and we again need to simplify the odd and even sums in (3.45). Observe that

$$(\Omega_e)^pY_d^j(\Omega_e)^q = \left((-1)^{2k-2}a_j^{2k-2}\right) Y_d^j = \left((-1)^{k-1}a_j^{2k-2}\right) Y_d^j$$

for $a_j$ the entry in the $j^{th}$ block along the diagonal of the equilibrium $\Omega_e$. Returning to the sum in (3.45), notice that there are $2k - 1$ distinct ways of choosing $p$ and $q$ such that $p + q = 2k - 2$. Likewise there are $2k - 1$ different ways of expressing the product $(\Omega_e)^pY_d^j(\Omega_e)^q$ with $p + q = 2k - 2$, and we therefore simplify the sum in (3.45),

$$\left. \frac{d}{dt} \right|_{t=0} (\Omega_e + tY_d^j)^{2k-1} = (2k - 1) \left((-1)^{k-1}a_j^{2k-2}\right) Y_d^j.$$ 

Finally, using the values of the trace form with respect to the basis $\{\tilde{Y}_j\}$ (see 3.27), we can compute each diagonal entry in the $n \times n$ block of the matrix $d^2C_k(\Omega_e)$:
\[
(d^2 C_k(\Omega_e))_{ij} = d^2 C_k(\Omega_e)(\tilde{Y}_i, Y^j_d)
\]
\[
= (-1)^{k-1} \left\langle \tilde{Y}_i, (2k - 1) \left( (-1)^{k-1} a_j^{2k-2} \right) Y^j_d \right\rangle
\]
\[
= (2k - 1)a_j^{2k-2} \left\langle \tilde{Y}_i, Y^j_d \right\rangle
\]
\[
= \begin{cases} 
(2k - 1)a_j^{2k-2} & \tilde{Y}_i = \tilde{Y}_j = Y^j_d \\
0 & \text{otherwise}
\end{cases}.
\tag{3.52}
\]

Therefore, setting \( \tilde{Y}_j = Y^j_d \) equation equation (3.52) implies that the \( j \)th column (for \( 1 \leq j \leq n \)) of \( d^2 C_k(\Omega_e) \) has all zero entries except for the entry \( (2k - 1)a_j^{2k-2} \) in row \( j \). Indeed the first \( n \times n \) block in \( d^2 C_k(\Omega_e) \) is of the form described in (3.3.4).

\section*{Theorem 3.3.6}

Let \( b \) denote the Pfaffian evaluated at \( \Omega_e \),

\[
b := Pf(\Omega_e) = a_1a_2 \ldots a_n.
\tag{3.53}
\]

The second variation of \( C_n(\Omega) = Pf(\Omega) \) at \( \Omega = \Omega_e \) can be represented by the \( (2n^2 - n) \times (2n^2 - n) \) symmetric, block diagonal matrix

\[
d^2 C_n(\Omega_e) = \begin{pmatrix}
P_{n \times n} & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & -P_{12} & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & 0 & P_{12} & \ldots & 0 & 0 & \ldots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ldots & -P_{r,c} & 0 & \ldots & 0 \\
0 & 0 & 0 & \ldots & 0 & P_{r,c} & \ldots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ldots & 0 & 0 & \ldots & P_{n-1,n}
\end{pmatrix}
\tag{3.54}
\]
where

$$P_n \times n = \begin{pmatrix}
0 & \frac{b}{a_1 a_2} & \frac{b}{a_1 a_3} & \cdots & \frac{b}{a_1 a_n} \\
\frac{b}{a_1 a_2} & 0 & \frac{b}{a_2 a_3} & \cdots & \frac{b}{a_2 a_n} \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\frac{b}{a_1 a_{n-1}} & \frac{b}{a_2 a_{n-1}} & \cdots & 0 & \frac{b}{a_{n-1} a_n} \\
\frac{b}{a_1 a_n} & \frac{b}{a_2 a_n} & \cdots & \frac{b}{a_{n-1} a_n} & 0
\end{pmatrix}$$  \((3.55)\)

and

$$P_{r,c} = \begin{pmatrix}
0 & \frac{b}{a_r a_c} \\
\frac{b}{a_r a_c} & 0
\end{pmatrix}.  \quad (3.56)$$

**Proof** The entries of the matrix $$d^2 C_n(\Omega_e)$$ are obtained as usual by

$$d^2 C_n(\Omega_e)(\tilde{Y}_i, \tilde{Y}_j) = \left\langle \tilde{Y}_i, \frac{d}{dt}igg|_{t=0} dC_n\left(\Omega_e + t\tilde{Y}_j\right) \right\rangle.  \quad (3.57)$$

Recall the derivative of Pf($$\Omega$$) derived in Lemma 3.2.3, $$dC_n(\Omega) = -\text{Pf}(\Omega) \cdot \Omega^{-1}$$. We now calculate the derivative of $$dC_n$$ at $$\Omega_e$$ using the derivative of the Pfaffian previously computed in (3.13)

$$\left. \frac{d}{dt} \right|_{t=0} dC_n\left(\Omega_e + t\tilde{Y}_j\right) = \left. \frac{d}{dt} \right|_{t=0} \left( -\text{Pf}(\Omega_e + t\tilde{Y}_j) \cdot (\Omega_e + t\tilde{Y}_j)^{-1} \right)$$

$$= \left. \frac{d}{dt} \right|_{t=0} \left( -\text{Pf}(\Omega_e + t\tilde{Y}_j) \cdot (\Omega_e)^{-1} - \text{Pf}(\Omega_e) \cdot \frac{d}{dt} \left|_{t=0} (\Omega_e + t\tilde{Y}_j)^{-1} \right) \right)$$

$$= \left[ -\frac{1}{2} \text{Pf}(\Omega_e) \cdot \text{tr}\left( \tilde{Y}_j(\Omega_e)^{-1} \right) \right] (\Omega_e)^{-1} - \text{Pf}(\Omega_e) \left( -\Omega_e^{-1} \tilde{Y}_j(\Omega_e)^{-1} \right)$$

$$= \kappa(\tilde{Y}_j)(\Omega_e)^{-1} - b \cdot K(\tilde{Y}_j)$$  \((3.58)\)

for $$b$$ in (3.53) and where we have defined the scalar $$\kappa(\tilde{Y}_j) = -\frac{1}{2} \text{Pf}(\Omega_e) \cdot \text{tr}(\tilde{Y}_j(\Omega_e)^{-1})$$ and the matrix $$K(\tilde{Y}_j) = -(\Omega_e)^{-1} \tilde{Y}_j(\Omega_e)^{-1}$$. We simplify both $$\kappa(\tilde{Y}_j)$$ and $$K(\tilde{Y}_j)$$ with respect to the basis $$\{\tilde{Y}_j\}$$,

$$\kappa(\tilde{Y}_j) = \begin{cases} 
-\frac{b}{a_j} & \tilde{Y}_j = Y_d^j \\
0 & \text{otherwise}
\end{cases}  \quad (3.59)$$
First assume \( \tilde{Y}_j = Y^j_d \) is a diagonal basis matrix (3.29). We substitute the values for \( \kappa(Y^j_d) \) and \( K(Y^j_d) \) from (3.59) and (3.60) respectively into the expression (3.58):

\[
K(\tilde{Y}_j) = \begin{pmatrix}
\frac{1}{a_j} Y_j & \tilde{Y}_j = Y^j_d \\
\frac{1}{a_j a_i} Y_{r,c}^{e22} & \tilde{Y}_j = Y^{e}_{r,c} \\
\frac{1}{a_j a_i} Y_{r,c}^{e11} & \tilde{Y}_j = Y^{e}_{r,c} \\
\frac{-1}{a_j a_i} Y_{r,c}^{e21} & \tilde{Y}_j = Y^{e}_{r,c} \\
\frac{-1}{a_j a_i} Y_{r,c}^{e12} & \tilde{Y}_j = Y^{e}_{r,c}
\end{pmatrix}
\] (3.60)

Notice in particular that when \( \tilde{Y}_i = \tilde{Y}_j = Y^j_d \), the sums \( \left(-\frac{b}{a_j}\right) \left(-\frac{1}{a_i}\right) \) and \( -\frac{b}{a_j^2} \cdot \delta_{ij} \) cancel, and the diagonal entries in first block vanish. Equation (3.61) yields the first \( n \) columns of the matrix \( d^2C_n(\Omega_e) \) and thus the entire block \( P_{n \times n} \).

Now assume \( \tilde{Y}_j \) in (3.57) is not a diagonal basis matrix. In particular, set \( \tilde{Y}_j = Y^{e_{11}}_{r,c} \). Again using the observations in (3.59) and (3.60) for the values of \( \kappa(Y^{e_{11}}_{r,c}) \) and \( K(Y^{e_{11}}_{r,c}) \), we simplify (3.58) in the calculation below:
\[
\left\langle \dot{Y}_i, \frac{d}{dt} \right|_{t=0} dC_n \left( \Omega_e + t\eta_{r,c}^{e11} \right) \right\rangle = \left\langle \dot{Y}_i, \kappa(Y_{r,c}^{e11}) \cdot (\Omega_e)^{-1} - b \cdot K(Y_{r,c}^{e11}) \right\rangle \\
= \left\langle \dot{Y}_i, 0 \cdot (\Omega_e)^{-1} - \frac{b}{a_r a_c} \cdot Y_{r,c}^{e22} \right\rangle \\
= -\frac{b}{a_r a_c} \left\langle \dot{Y}_i, Y_{r,c}^{e22} \right\rangle \\
= \left\{ \begin{array}{ll}
-\frac{b}{a_r a_c} \dot{Y}_i = Y_{r,c}^{e22} \\
0 & \dot{Y}_i \neq Y_{r,c}^{e22}
\end{array} \right. \\
(3.62)
\]

By symmetry of the form \langle , \rangle in (3.6), we have \(d^2C_n(\Omega_e)(\dot{Y}_i, \dot{Y}_j) = d^2C_n(\Omega_e)(\dot{Y}_j, \dot{Y}_i)\), we have \(d^2C_n(\Omega_e)(Y_{r,c}^{e11}, Y_{r,c}^{e22}) = d^2C_n(\Omega_e)(Y_{r,c}^{e22}, Y_{r,c}^{e11})\), and we obtain the smaller block \(-P_{r,c}\) along the diagonal.

Similarly, due to the leading coefficient \(-1\) corresponding to both \(K(Y_{r,c}^{e12})\) and \(K(Y_{r,c}^{e21})\) in (3.60), the pairing \(d^2C_n(\Omega_e)(Y_{r,c}^{e12}, Y_{r,c}^{e21}) = d^2C_n(\Omega_e)(Y_{r,c}^{e21}, Y_{r,c}^{e12})\) (and zero otherwise). Therefore the blocks \(P_{r,c}\) indeed appear along the diagonal of \(d^2C_n(\Omega_e)\), and moreover these blocks appear immediately after each matrix \(-P_{r,c}\).

For each \(M_{r,c}\) we get a pair of consecutive 2 \(\times\) 2 matrices \(-P_{r,c}\) and \(P_{r,c}\) along the diagonal, and thus the proof of Theorem 3.3.6 is complete.

Since we have now shown that matrices \(d^2H\) and \(d^2C_k\) (for \(1 \leq k \leq n\)) are all block diagonal, we obtain a block diagonal matrix for the terms \(d^2H(\Omega_e) + \text{First}(\Omega_e)\) from Proposition 3.3.1 which is summarized in the corollary below:

**Corollary 3.3.7** For \(1 \leq k \leq n\), let \(\phi_k = \phi_k(F(\Omega_e))\) denote the entries of the column vector \(\Phi'\) defined in Theorem 3.2.4. The sum \(d^2H(\Omega_e) + \text{First}(\Omega_e)\) can be expressed as a block diagonal, symmetric matrix.
\[ V = d^2 H(\Omega_e) + \sum_{k=1}^{n} (\phi_k \cdot d^2 C_k(\Omega_e)) \]

\[
\left( \begin{array}{cccccc}
V_{n \times n} & 0 & 0 & \ldots & 0 & 0 \\
0 & V_{1,2}^{\text{diag}} & 0 & \ldots & 0 & 0 \\
0 & 0 & V_{1,2}^{\text{off}} & \ldots & 0 & 0 \\
0 & 0 & 0 & \ldots & V_{n-1,n}^{\text{diag}} & 0 \\
0 & 0 & 0 & \ldots & 0 & V_{n-1,n}^{\text{off}} \\
\end{array} \right).
\] (3.63)

The \( n \times n \) block \( V_{n \times n} \) resulting from the relations among the diagonal basis elements \((Y_i, Y_j) = (Y_d^i, Y_d^j)\) is defined by

\[
(V_{n \times n})_{ij} = \begin{cases} 
(\lambda_{2i-1} + \lambda_{2i}) + \sum_{k=1}^{n-1} [(\phi_k)(2k-1)(a_1^{2k-2})] & 1 \leq i = j \leq n \\
(\phi_n) \frac{b}{a_i a_j} & 1 \leq i \neq j \leq n
\end{cases}.
\]

For each of the \( \frac{n^2-n}{2} \) blocks \( M_{r,c} \) spanned by the basis elements \( Y_{r,c}^{e11}, Y_{r,c}^{e22}, Y_{r,c}^{e12}, \)

and \( Y_{r,c}^{e21} \), the two \( 2 \times 2 \) blocks \( V_{r,c}^{\text{diag}} \) and \( V_{r,c}^{\text{off}} \) resulting from the diagonal pairings \((Y_{r,c}^{e11}, Y_{r,c}^{e22})\) and anti-diagonal pairings \((Y_{r,c}^{e12}, Y_{r,c}^{e21})\), respectively, are defined by

\[
V_{r,c}^{\text{diag}} = \begin{pmatrix} (\lambda_{2r-1} + \lambda_{2r}) + \sum_{k=1}^{n-1} [(\phi_k)(E_{2k-2}(a_r, a_c))] \\ \sum_{k=1}^{n-1} [(\phi_k)(E_{2k-2}(a_r, a_c))] \end{pmatrix} - \begin{pmatrix} \sum_{k=1}^{n-1} [(\phi_k)(O_{2k-2}(a_r, a_c))] \end{pmatrix} - \begin{pmatrix} \phi_n \left( \frac{b_{r,c}}{a_r a_c} \right) \end{pmatrix}
\]

and

\[
V_{r,c}^{\text{off}} = \begin{pmatrix} (\lambda_{2r-1} + \lambda_{2r}) + \sum_{k=1}^{n-1} [(\phi_k)(E_{2k-2}(a_r, a_c))] \\ \sum_{k=1}^{n-1} [(\phi_k)(E_{2k-2}(a_r, a_c))] \end{pmatrix} - \begin{pmatrix} \sum_{k=1}^{n-1} [(\phi_k)(O_{2k-2}(a_r, a_c))] \end{pmatrix} + \begin{pmatrix} \phi_n \left( \frac{b_{r,c}}{a_r a_c} \right) \end{pmatrix}
\]
3.3.4 Constructing the Matrix of Second Partials

Lastly it remains to construct a matrix for the final piece of equation (3.21), namely the sum $\text{Second}(\Omega_e)$ in (3.22) of terms involving the second partial derivatives $\phi_{kl}$.

**Theorem 3.3.8** Using the basis $\{Y_i\}$, the sum of terms in $\text{Second}(\Omega_e)$ (3.22) can be represented by the $(2n^2 - n) \times (2n^2 - n)$ symmetric, block diagonal matrix

$$
S = \begin{pmatrix}
S_{n \times n} & 0 & \cdots & 0 \\
0 & \ddots & & \\
\vdots & & \ddots & \\
0 & & & 0
\end{pmatrix},
$$

where the lone non-trivial $n \times n$ block $S_{n \times n}$ is given by

$$
S_{n \times n} = d\tilde{F}_{\Omega_e}^t \text{Hess}(\phi) d\tilde{F}_{\Omega_e}.
$$

Again, $d\tilde{F}_{\Omega_e}$ denotes the Jacobian of the restriction of $F|_b = \tilde{F}$ at $\Omega_e$ previously introduced in Theorem 3.2.4, and $\text{Hess}(\phi)$ denotes the $n \times n$ Hessian of $\phi$ at $\Omega_e$.

**Proof** For $1 \leq k \leq n$, recall from (3.11) and (3.14) that $dC_k(\Omega_e)$ is a block diagonal matrix. Thus if either $\tilde{Y}_i$ or $\tilde{Y}_j$ is set as an off-diagonal basis element $Y^{r,ij}_{r,\tilde{c}}$, then at least one of the inner products $\langle \tilde{Y}_i, dC_k(\Omega_e) \rangle$ or $\langle \tilde{Y}_j, dC_k(\Omega_e) \rangle$ in the expression for $\text{Second}(\Omega_e)$ (see 3.22)) will vanish for all $k$. In such a case the entire sum $\text{Second}(\Omega_e)$ is trivial. Equivalently, the only non-zero entries in $S$ appear in the first $n \times n$ block, denoted $S_{n \times n}$. We now focus our attention solely on the case when $(\tilde{Y}_i, \tilde{Y}_j) = (Y^i_d, Y^j_d)$ are pairs of diagonal matrices.

For a given diagonal matrix $Y^j_d$, equations (3.11) and (3.14, together with the
trace relations given in (3.27), imply
\[
\langle Y^j_d, dC_k(\Omega_e) \rangle = \begin{cases} 
   a^{2k-1}_j & \text{if } 1 \leq k \leq n - 1 \\
   \frac{b}{a_j} & \text{if } k = n
\end{cases} 
\]  
(3.66)

with \( b \) the constant defined in (3.53). Let \( \phi_{kl} \) denote the value of \( \phi_{kl} \) at the point \( F(\Omega_e) \in \mathbb{R}^n \). Using the relations in (3.66) above, we simplify \( \text{Second}(\Omega_e) \) for \((\tilde{Y}_i, \tilde{Y}_j) = (Y^i_d, Y^j_d)\)

\[
(S_{n\times n})_{ij} = \sum_{k=1}^{n} \left\{ \langle Y^i_d, dC_k(\Omega_e) \rangle \sum_{l=1}^{n} \phi_{kl} \langle Y^j_d, dC_l(\Omega_e) \rangle \right\} 
= I_1 + I_2 + I_3 + I_4 
\]  
(3.67)

where the sum has been separated into four parts:

\[
I_1 = \langle Y^j_d, dC_n(\Omega_e) \rangle \cdot \sum_{k=1}^{n-1} \left\{ \phi_{kn} \cdot \langle Y^i_d, dC_k(\Omega_e) \rangle \right\} \quad \text{when } k \neq n, \ l = n \\
I_2 = \langle Y^i_d, dC_n(\Omega_e) \rangle \cdot \sum_{l=1}^{n-1} \left\{ \phi_{ln} \cdot \langle Y^j_d, dC_l(\Omega_e) \rangle \right\} \quad \text{when } k = n, \ l \neq n \\
I_3 = \sum_{k=1}^{n-1} \left\{ \langle Y^i_d, dC_k(\Omega_e) \rangle \cdot \sum_{l=1}^{n-1} \left[ \phi_{kl} \cdot \langle Y^j_d, dC_l(\Omega_e) \rangle \right]\right\} \quad \text{when } k \neq n, \ l \neq n \\
I_4 = \phi_{nn} \cdot \langle Y^i_d, dC_n(\Omega_e) \rangle \langle Y^j_d, dC_n(\Omega_e) \rangle \quad \text{when } k = l = n.
\]

It is convenient to separate the sum in this manner since \( k = n \) is the special case corresponding to the Pfaffian Casimir \( dC_n(\Omega) \). Equations (3.66) and (3.27) therefore
\[ I_1 = \sum_{k=1}^{n-1} \left[ \left( \frac{b}{a_j} \right) \left( a_i^{2k-1} \right) \phi_{kn} \right] \]
\[ I_2 = \sum_{l=1}^{n-1} \left[ \left( \frac{b}{a_i} \right) \left( a_j^{2l-1} \right) \phi_{ln} \right] \]
\[ I_3 = \sum_{l=1}^{n-1} \left\{ \left( a_j^{2l-1} \right) \sum_{k=1}^{n-1} \left[ (\phi_{kl}) \left( a_i^{2k-1} \right) \right] \right\} \]
\[ I_4 = \left( \frac{b}{a_j} \right) \left( \frac{b}{a_i} \right) \phi_{nn}. \]

On the other hand consider, the product \( d\widetilde{F}_{\Omega_e}^t \text{Hess}(\phi) \, d\widetilde{F}_{\Omega_e} \), whose \( ij \)th entry is computed in Appendix B:

\[
\left( d\widetilde{F}_{\Omega_e}^t \text{Hess}(\phi) \, d\widetilde{F}_{\Omega_e} \right)_{ij} = \left( \begin{array}{cccc}
    a_i & \phi_{11} & \phi_{12} & \phi_{13} & \cdots & \phi_{1n} \\
    a_i^3 & \phi_{12} & \phi_{22} & \phi_{23} & \cdots & \phi_{2n} \\
      \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
    a_i^{2n-3} & \phi_{1,n-1} & \phi_{2,n-1} & \phi_{3,n-1} & \cdots & \phi_{n-1,n} \\
    \frac{b}{a_j} & \phi_{1n} & \phi_{2n} & \phi_{3n} & \cdots & \phi_{nn}
\end{array} \right) \left( \begin{array}{c}
    a_j \\
    a_j^3 \\
      \vdots \\
    a_j^{2n-3} \\
    \frac{b}{a_j}
\end{array} \right) \\
= \sum_{l=1}^{n-1} \left[ a_j^{2l-1} \sum_{k=1}^{n-1} \phi_{lk} \right] + \sum_{l=1}^{n-1} \left[ a_j^{2l-1} \frac{b}{a_j} \phi_{ln} \right] \\
+ \sum_{k=1}^{n-1} \left( \frac{b}{a_i} a_j^{2k-1} \phi_{kn} \right) + \frac{b^2}{a_i a_j} \phi_{nn} \\
= I_1 + I_2 + I_3 + I_4. 
\]

Thus, we have shown

\[ (S_{n\times n})_{ij} = I_1 + I_2 + I_3 + I_4 = \left( d\widetilde{F}_{\Omega_e}^t \text{Hess}(\phi) \, d\widetilde{F}_{\Omega_e} \right)_{ij}. \]

As a result of Corollary 3.3.7 and Theorem 3.3.8 we have indeed expressed \( d^2 H_\phi(\Omega_e) \) as the sum of \((2n^2 - n) \times (2n^2 - n)\) symmetric, block diagonal matrices
V and S, which we summarize in the following corollary:

**Corollary 3.3.9** Using the basis \{\hat{Y}_i\}, the second variation of \(H_\phi\) at a generic equilibrium \(\Omega_e\) (see 2.19) is represented by the \((2n^2 - n) \times (2n^2 - n)\) symmetric, block diagonal matrix

\[
d^2H_\phi(\Omega_e) = V + S
\]

\[
\begin{pmatrix}
V_n \times n + S_n \times n & 0 & 0 & \ldots & 0 & 0 \\
0 & V^{\text{diag}}_{1,2} & 0 & \ldots & 0 & 0 \\
0 & 0 & V^{\text{off}}_{1,2} & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & V^{\text{diag}}_{n-1,n} & 0 \\
0 & 0 & 0 & \ldots & 0 & V^{\text{off}}_{n-1,n}
\end{pmatrix}
\]

where \(V = d^2H(\Omega_e) + \text{First}(\Omega_e)\) is defined in Corollary 3.3.7 and \(S\) is the matrix in the previous Theorem 3.3.8.

Corollary 3.3.9 is a significant result in ultimately studying the definiteness of the second variation. Definiteness of the second variation is now equivalent to definiteness of each block in \(d^2H_\phi(\Omega_e)\). Before analyzing the definiteness of the blocks in (3.68), we pause to make several important observations regarding the final representation of \(d^2H_\phi(\Omega_e)\) in (3.68):

- We denote the first \(n \times n\) block

\[
T_{n \times n} = V_{n \times n} + S_{n \times n}.
\]

This block is used to evaluate the second variation \(d^2H_\phi(\Omega_e)\) at a pair of diagonal elements \(Y\) and \(Z\) in the Cartan subalgebra \(\mathfrak{h}\) defined in (2.4).

- The matrix \(V_{n \times n} = d^2H(\Omega_e) + \text{First}(\Omega_e)\) is independent of the second partial derivatives \(\phi_{kl}\) at \(F(\Omega_e)\). The entries of \(V_{n \times n}\) are expressions involving the \(\lambda_i\)'s
from the inertia matrix $J$ (see 1.36), the $a_i$’s from the generic equilibrium $\Omega_e$ (see 2.19), and the first partial derivatives $\phi_k$ at $F(\Omega_e)$ obtained in Theorem 3.2.4.

- The preceding comment implies that the second partial derivatives $\phi_{kl}(F(\Omega_e))$ only appear in the matrix $S_{n \times n} = \tilde{d}F^t_{\Omega_e} \text{Hess}(\phi) \tilde{d}F_{\Omega_e}$.

- Relations among the entries of each $2 \times 2$ block $M_{r,e}$ pictured in (3.28) result in two neighboring $2 \times 2$ blocks $V_{r,e}^{\text{diag}}$ followed by $V_{r,e}^{\text{off}}$, both defined in Corollary 3.3.7. These $2 \times 2$ blocks are independent of $\phi_{kl}(F(\Omega_e))$.

- From the block diagonal structure of the matrix $d^2H_{\phi}(\Omega_e)$ in (3.68), it is clear that pairing any element $Y$ in the Cartan subalgebra $\mathfrak{h}$ with any element $Z$ from the rest of the Lie algebra $\mathfrak{so}(2n)$ yields a trivial result:

$$d^2H_{\phi}(\Omega_e) \langle Y, Z \rangle = 0 \text{ for } Y \in \mathfrak{h} \text{ and } Z \notin \mathfrak{h}$$

### 3.4 Definiteness of the Second Variation

Thus far we have shown in Theorem 3.2.4 it is possible to construct a function $\phi$ such that the family of Casimirs $\phi_F(\Omega)$ in (3.4) has a critical point at $\Omega_e$. Namely if the first partial derivatives of $\phi$ at $F(\Omega_e)$ satisfy equation (3.16), then a generic equilibrium $\Omega_e$ is a critical point of $H_{\phi}$. We now wish to find conditions such that the second variation of $H_{\phi}$ is definite (either positive or negative) at $\Omega_e$. Then, by the energy-Casimir method, $\Omega_e$ is a stable equilibrium of the Hamiltonian system $X_H = [J, \Omega^2]$ (see 2.6). In this section we will first find necessary and sufficient conditions on the inertia matrix $J$ in (1.36) such that all of the $2 \times 2$ blocks in the matrix $d^2H_{\phi}(\Omega_e)$ given in Corollary 3.3.9 are simultaneously positive or simultaneously negative definite. After working with the $2 \times 2$ blocks we will turn our attention to the $n \times n$ block $T_{n \times n} = V_{n \times n} + S_{n \times n}$ of the second variation matrix $d^2H_{\phi}(\Omega_e)$ and
find further sufficient conditions on the arbitrary function \( \phi \) which will ensure that 
\[ d^2 H_\phi(\Omega_c) \] is definite.

### 3.4.1 Definiteness of the 2 × 2 Blocks

Recall from Corollary 3.3.7 that there are \((n^2 - n) \) 2 × 2 blocks along the diagonal of 
\[ d^2 H_\phi(\Omega_c) \] which come in consecutive pairs \( V^\text{diag}_{r,c} \) and \( V^\text{off}_{r,c} \). In general, a 2 × 2 matrix \( M \) is positive definite if and only if both eigenvalues of \( M \) are positive, and negative definite if and only if both eigenvalues are negative. Consequently, if both the trace and determinant of the 2 × 2 matrix are positive, the matrix is positive definite. Similarly, if the determinant is positive but the trace is negative, the matrix is negative definite. We will now compute formulas for the trace and determinant of the blocks \( V^\text{diag}_{r,c} \) and \( V^\text{off}_{r,c} \) in order to find conditions on their signs.

Before computing the trace and determinants of these blocks, we first introduce the following vectors of even and odd sums, \( E_p(a_r, a_c) \) and \( O_p(a_r, a_c) \), defined in (3.39):

\[
E = \begin{pmatrix}
E_0(a_r, a_c) \\
E_2(a_r, a_c) \\
E_4(a_r, a_c) \\
\vdots \\
E_{2n-4}(a_r, a_c) \\
0
\end{pmatrix} = \begin{pmatrix}
1 \\
\frac{a_r^2 + a_c^2}{2} \\
\frac{a_r^4 + a_r^2 a_c^2 + a_c^4}{4} \\
\vdots \\
\frac{a_r^{2n-4} + a_r^{2n-6} a_c^2 + \ldots + a_c^{2n-4}}{2n-4}
\end{pmatrix}
\]

(3.70)

\[
O = \begin{pmatrix}
O_0(a_r, a_c) \\
O_2(a_r, a_c) \\
O_4(a_r, a_c) \\
\vdots \\
O_{2n-4}(a_r, a_c) \\
-\prod_{i \neq r,c} a_i
\end{pmatrix} = \begin{pmatrix}
0 \\
\frac{a_r a_c}{2} \\
\frac{a_r^3 a_c + a_r a_c^3}{3} \\
\vdots \\
\frac{a_r^{2n-5} a_c + a_r^{2n-7} a_c^3 + \ldots + a_r a_c^{2n-5}}{2n-5}
\end{pmatrix}
\]

(3.71)

with \( b \) the value of \( Pf(\Omega_c) \) defined in (3.53). Using these two vectors (with \( n \)
entries) we give a new expression for the matrices $V_{r,c}^{\text{diag}}$ and $V_{r,c}^{\text{off}}$ of Corollary 3.3.7:

\[
V_{r,c}^{\text{diag}} = \begin{pmatrix}
(\lambda_{2r-1} + \lambda_{2c-1}) + \langle \Phi', \hat{E} \rangle & \langle \Phi', \hat{O} \rangle \\
\langle \Phi', \hat{O} \rangle & (\lambda_{2r} + \lambda_{2c}) + \langle \Phi', \hat{E} \rangle
\end{pmatrix} \tag{3.72}
\]

\[
V_{r,c}^{\text{off}} = \begin{pmatrix}
(\lambda_{2r-1} + \lambda_{2c}) + \langle \Phi', \hat{E} \rangle & -\langle \Phi', \hat{O} \rangle \\
-\langle \Phi', \hat{O} \rangle & (\lambda_{2r} + \lambda_{2c-1}) + \langle \Phi', \hat{E} \rangle
\end{pmatrix} \tag{3.73}
\]

where $\langle \ , \ \rangle$ denotes the standard Euclidean product.

These two matrices are almost identical, and notice in particular that $\tr V_{r,c}^{\text{diag}} = \tr V_{r,c}^{\text{off}}$. The last entries 0 and $-\frac{b}{a_r a_c}$ in the vectors $\hat{E}$ and $\hat{O}$, respectively, are distinct from the other entries in that they are not even or odd sums such as (3.39). Recall that the last entry in the vector $\Phi'$ of partial derivatives is $\phi_n(F(\Omega_c))$, the first partial derivative in the direction of the Pfaffian Casimir $C_n(\Omega)$. Therefore, when they are paired with $\Phi'$, it is reasonable to expect the last entries of $\hat{E}$ and $\hat{O}$ to be somewhat special.

We now find a formula for the trace of the matrices $V_{r,c}^{\text{off}}$ and $V_{r,c}^{\text{off}}$. We begin with the lemma below which will be a crucial component in calculating the trace.

**Lemma 3.4.1** Consider the following vector $\gamma \in \mathbb{R}^n$ with the following entries,

\[
\gamma_i = \begin{cases}
\frac{a_i^2}{a_r^2 - a_c^2} & i = r \\
\frac{-a_i^2}{a_r^2 - a_c^2} & i = c \\
0 & 1 \leq i \leq n, \ i \neq r, c
\end{cases} \tag{3.74}
\]

For $d\tilde{F}_{\Omega_c}$ and $A = \text{diag}(a_1, a_2, \ldots, a_n)$ defined in Theorem 3.2.4, it follows that

\[
d\tilde{F}_{\Omega_c} A^{-1} \gamma = \hat{E}. \tag{3.75}
\]

**Proof** Substituting $A^{-1} = \text{diag} \left( \frac{1}{a_1}, \ldots, \frac{1}{a_n} \right)$ into the left side of equation (3.75) we
Lemma 3.4.1 equivalently states that \( \gamma \) is equal to the product

\[
A \ d\tilde{F}_{\Omega_c}^{-1} \ E = \gamma.
\]  

Equation (3.76) is important in the calculation of the trace summarized in the theorem below:

**Theorem 3.4.2** For all \( 1 \leq r < c \leq n \), the trace of the \( 2 \times 2 \) blocks can be simplified
as follows,

\[ \text{tr} \left( V_{r,c}^{\text{diag}} \right) = \text{tr} \left( V_{r,c}^{\text{off}} \right) = \frac{a_r^2 + a_c^2}{a_r^2 - a_c^2} \left( \lambda_{2c-1} + \lambda_{2c} - \lambda_{2r-1} - \lambda_{2r} \right). \quad (3.77) \]

**Proof** From the general form of \( V_{r,c}^{\text{diag}} \) and \( V_{r,c}^{\text{off}} \) in (3.72) and (3.73) respectively,

\[ \text{tr} \left( V_{r,c}^{\text{diag}} \right) = \text{tr} \left( V_{r,c}^{\text{off}} \right) = \left( \lambda_{2r-1} + \lambda_{2r} + \lambda_{2c-1} + \lambda_{2c} \right) + 2 \ll \Phi', \widehat{E} \gg. \]

The inner product \( \ll \Phi', \widehat{E} \gg \) can be computed by first using Theorem 3.2.4 to replace \( \Phi' \),

\[ \ll \Phi', \widehat{E} \gg = \ll - \left( dF_{\Omega_e}^t \right)^{-1} A \Lambda, \widehat{E} \gg \]

We move the matrices \( (dF_{\Omega_e}^t)^{-1} \) and \( A \) through the inner product above by transposing their product (recall \( A \) is diagonal so \( A^t = A \)) and simplify using equation (3.76),

\[ \ll - \left( dF_{\Omega_e}^t \right)^{-1} A \Lambda, \widehat{E} \gg = \ll \Lambda, -A dF_{\Omega_e}^{-1} \widehat{E} \gg \\
\ll \Lambda, -\gamma \gg \\
= -a_r^2 (\lambda_{2r-1} + \lambda_{2r}) + a_c^2 (\lambda_{2c-1} + \lambda_{2c}) \frac{a_r^2 - a_c^2}{a_r^2 - a_c^2}. \quad (3.78) \]

Finally, substituting equation (3.78) into the calculation of the trace, we simplify
further:

\[ \begin{align*}
&\text{tr} (V_{r,c}) = (\lambda_{2r-1} + \lambda_{2r} + \lambda_{2c-1} + \lambda_{2c}) + 2 \ll \Phi', \hat{E} \gg \\
&\quad = (\lambda_{2r-1} + \lambda_{2r} + \lambda_{2c-1} + \lambda_{2c}) + 2 \left( \frac{-a_r^2 (\lambda_{2r-1} + \lambda_{2r}) + a_c^2 (\lambda_{2c-1} + \lambda_{2c})}{a_r^2 - a_c^2} \right) \\
&\quad = a_r^2 (\lambda_{2r-1} + \lambda_{2r}) - a_c^2 (\lambda_{2c-1} + \lambda_{2c}) + a_r^2 (\lambda_{2c-1} + \lambda_{2c}) - a_c^2 (\lambda_{2c-1} + \lambda_{2c}) \\
&\quad \quad + \frac{-2a_r^2 (\lambda_{2r-1} + \lambda_{2r}) + 2a_c^2 (\lambda_{2c-1} + \lambda_{2c})}{a_r^2 - a_c^2} \\
&\quad = a_r^2 + a_c^2 \left( \lambda_{2c-1} + \lambda_{2c} - \lambda_{2r-1} - \lambda_{2r} \right). 
\end{align*} \]

The trace in (3.77) holds for each 2 × 2 block along the diagonal. Recall in Section 2.4 we have assumed, without any loss of generality, that the entries of a generic equilibrium \( \Omega_e \) defined in (2.19) are ordered such that \(|a_r| < |a_c|\) when \( r < c \). Thus the coefficient \( \frac{a_r^2 + a_c^2}{a_r^2 - a_c^2} \) calculated in Theorem 3.4.2 must be negative. Based on this observation and Theorem 3.4.2, the corollary below provides necessary and sufficient conditions for the trace of a given 2 × 2 block \( V_{r,c}^{diag} \) or \( V_{r,c}^{off} \) to be positive or negative.

**Corollary 3.4.3** For \( 1 \leq r < c \leq n \), consecutive blocks \( V_{r,c}^{diag} \) and \( V_{r,c}^{off} \) both have positive trace if and only if the \( \lambda_i \)'s of the inertia matrix \( J \) in (1.36) satisfy the inequality

\[ \lambda_{2r-1} + \lambda_{2r} > \lambda_{2c-1} + \lambda_{2c}. \]  \hspace{1cm} (3.79)

Likewise, the traces of \( V_{r,c}^{diag} \) and \( V_{r,c}^{off} \) are both negative if and only if the \( \lambda_i \)'s satisfy

\[ \lambda_{2r-1} + \lambda_{2r} < \lambda_{2c-1} + \lambda_{2c}. \]  \hspace{1cm} (3.80)

The traces of all pairs \( V_{r,c}^{diag} \) and \( V_{r,c}^{off} \) corresponding to each block \( M_{r,c} \) in the upper triangle must satisfy condition (3.79) or (3.80). In order for the traces of all blocks to be simultaneously positive we must satisfy \((n^2 - n)\) inequalities of the form (3.79). Likewise, for all of the 2 × 2 blocks to have negative traces, \((n^2 - n)\) inequalities of
the form (3.80) must be satisfied for all pairs \( r \) and \( c \) with \( 1 \leq r < c \leq n \).

Simplification of the determinants of the blocks \( V^{\text{diag}}_{r,c} \) and \( V^{\text{off}}_{r,c} \) can be performed in a similar fashion. First we prove a lemma similar to Lemma 3.4.1 in order to ultimately simplify the product \( \ll \Phi', \hat{O} \gg \).

**Lemma 3.4.4** Consider the following vector \( \eta \in \mathbb{R}^n \) with entries

\[
\eta_i = \begin{cases} \frac{a_r a_c}{a_r^2 - a_c^2} & i = r \\ \frac{-a_r a_c}{a_r^2 - a_c^2} & i = c \\ 0 & 1 \leq i \leq n, \ i \neq r,c \end{cases}, \quad (3.81)
\]

For \( d\tilde{F}\Omega_e \) and \( A = \text{diag}(a_1, a_2, \ldots, a_n) \) in Theorem 3.2.4, the following equation is satisfied:

\[
d\tilde{F}\Omega_e A^{-1} \eta = \hat{O}. \quad (3.82)
\]
Proof We proceed as in the proof of Lemma 3.4.1:

\[
\begin{pmatrix}
0 \\
\vdots \\
\frac{a_r}{a_r^2 - a_c} \\
\vdots \\
\frac{-a_r}{a_r^2 - a_c} \\
0
\end{pmatrix}
\]

\[
d\tilde{F}_{\Omega_e}^{-1} \eta = d\tilde{F}_{\Omega_e} \frac{1}{a_r^2 - a_c^2}
\begin{pmatrix}
a_1 & a_2 & \ldots & a_n \\
a_1^3 & a_2^3 & \ldots & a_n^3 \\
\vdots & \vdots & \ddots & \vdots \\
a_1^{2n-3} & a_2^{2n-3} & \ldots & a_n^{2n-3} \\
\frac{b}{a_1} & \frac{b}{a_2} & \ldots & \frac{b}{a_n}
\end{pmatrix}
\begin{pmatrix}
0 \\
\vdots \\
a_c \\
\vdots \\
-a_r \\
\vdots \\
0
\end{pmatrix}
\]

\[
= \frac{1}{a_r^2 - a_c^2}
\begin{pmatrix}
a_r a_c - a_r a_c \\
\vdots \\
a_r^3 a_c - a_r a_c^3 \\
\vdots \\
a_r^{2n-3} a_c - a_r a_c^{2n-3} \\
(a_r^2 - a_c^2) \left( -\frac{b}{a_r a_c} \right)
\end{pmatrix}
\]

\[
= \hat{O}.
\]

As a result of Lemma 3.4.4 we use the equivalent expression

\[
A d\tilde{F}_{\Omega_e}^{-1} \hat{O} = \eta.
\]  

(3.83)
to simplify the dot product $\ll \Phi', \hat{O} \gg$:

$$\ll \Phi', \hat{O} \gg = \ll - \left( dF_{\Omega_c}^t \right)^{-1} A \Lambda, \hat{O} \gg$$

$$= \ll \Lambda, -A \, dF_{\Omega_c}^{-1} \hat{O} \gg$$

$$= \ll \Lambda, -\eta \gg$$

$$= -a_r a_c (\lambda_{2r-1} + \lambda_{2r}) + a_r a_c (\lambda_{2c-1} + \lambda_{2c}) \over a_r^2 - a_c^2.$$

(3.84)

Using the simplification of the inner product above, we are now prepared to simplify the determinants of $V_{r,c}^{diag}$ and $V_{r,c}^{off}$. Notice from equations (3.72) and (3.73) that the determinants of $V_{r,c}^{diag}$ and $V_{r,c}^{off}$ are not the same, and we must handle each separately. First the determinant of $V_{r,c}^{diag}$ is simplified in the theorem below:

**Theorem 3.4.5** For all $1 \leq r < c \leq n,$

$$\text{det} (V_{r,c}^{diag}) = (\lambda_{2c} - \lambda_{2r-1})(\lambda_{2c-1} - \lambda_{2r}).$$

(3.85)

**Proof** Recall the general form of $V_{r,c}^{diag}$ from (3.72), and let $v_{11}$ and $v_{22}$ denote the two diagonal entries of $V_{r,c}^{diag}$ which we simplify using formula (3.78) for the product $\ll \Phi', \hat{E} \gg$:

$$v_{11} = (\lambda_{2r-1} + \lambda_{2c-1}) + \ll \Phi', \hat{E} \gg$$

$$= (a_r^2 - a_c^2)(\lambda_{2r-1} + \lambda_{2c-1}) \over a_r^2 - a_c^2$$

$$= (\lambda_{2c} - \lambda_{2r-1})a_c^2 + (\lambda_{2c-1} - \lambda_{2r})a_r^2 \over a_r^2 - a_c^2.$$

(3.86)

and likewise

$$v_{22} = (\lambda_{2r} + \lambda_{2c}) + \ll \Phi', \hat{E} \gg=$$

$$= (\lambda_{2c} - \lambda_{2r-1})a_r^2 + (\lambda_{2c-1} - \lambda_{2r})a_c^2 \over a_r^2 - a_c^2.$$

(3.87)

Next we simplify the product of the diagonal entries $v_{11}$ and $v_{22}$ using equations
(3.86) and (3.87),

\[v_{11}v_{22} = \left(\lambda_{2r-1} + \lambda_{2r-1}\right) + \left\langle \Phi', \hat{E} \right\rangle \left(\lambda_{2r} + \lambda_{2c}\right) + \left\langle \Phi', \hat{E} \right\rangle\]

\[= \left(\frac{(\lambda_{2c} - \lambda_{2r-1})a_c^2 + (\lambda_{2c-1} - \lambda_{2r})a_r^2}{a_r^2 - a_c^2}\right) \left(\frac{(\lambda_{2c} - \lambda_{2r-1})a_c^2 + (\lambda_{2c-1} - \lambda_{2r})a_r^2}{a_r^2 - a_c^2}\right)\]

\[= \left[\left(\lambda_{2c} - \lambda_{2r-1}\right)^2 + (\lambda_{2c-1} - \lambda_{2r})^2\right] a_r^2 a_c^2 + (\lambda_{2c} - \lambda_{2r-1})(\lambda_{2c-1} - \lambda_{2r})(a_r^4 + a_c^4)\]

\[(a_r^2 - a_c^2)^2.\]  

Likewise we use formula (3.84) for the product of the anti-diagonal entries \(\left\langle \Phi', \tilde{O} \right\rangle\) of \(V^\text{diag}_{r,c}\) to simplify

\[ \left[\left\langle \Phi', \tilde{O} \right\rangle\right]^2 = \left(\frac{-a_r a_c (\lambda_{2r-1} + \lambda_{2r}) + a_r a_c (\lambda_{2c-1} + \lambda_{2c})}{a_r^2 - a_c^2}\right)^2\]

\[= \left(\frac{(\lambda_{2c} - \lambda_{2r-1})a_r a_c + (\lambda_{2c-1} - \lambda_{2r})a_r a_c}{a_r^2 - a_c^2}\right)^2\]

\[= \left[\left(\lambda_{2c} - \lambda_{2r-1}\right)^2 + (\lambda_{2c-1} - \lambda_{2r})^2 + 2(\lambda_{2c} - \lambda_{2r-1})(\lambda_{2c-1} - \lambda_{2r})\right] a_r^2 a_c^2\]

\[(a_r^2 - a_c^2)^2.\]  

Notice that several terms in (3.88) and (3.89) cancel immediately, and thus the determinant of \(V^\text{diag}_{r,c}\) simplifies nicely:

\[\det\left(V^\text{diag}_{r,c}\right) = v_{11}v_{22} - \left[\left\langle \Phi', \tilde{O} \right\rangle\right]^2\]

\[= (\lambda_{2c} - \lambda_{2r-1})(\lambda_{2c-1} - \lambda_{2r})(a_r^4 + a_c^4) - 2(\lambda_{2c} - \lambda_{2r-1})(\lambda_{2c-1} - \lambda_{2r})a_r^2 a_c^2\]

\[= \frac{\left(a_r^4 - 2a_r^2 a_c^2 + a_c^4\right)}{(a_r^2 - a_c^2)^2}\]

\[\left(\lambda_{2c} - \lambda_{2r-1}\right)\left(\lambda_{2c-1} - \lambda_{2r}\right)\]

\[= (\lambda_{2c} - \lambda_{2r-1})(\lambda_{2c-1} - \lambda_{2r}).\]  

The matrix \(V^\text{off}_{r,c}\) in (3.73) is almost identical to \(V^\text{diag}_{r,c}\), and the determinant of \(V^\text{off}_{r,c}\) can be expressed as a similar product.

**Theorem 3.4.6** For all \(1 \leq r < c \leq n\),

\[\det\left(V^\text{off}_{r,c}\right) = (\lambda_{2c} - \lambda_{2r})(\lambda_{2c-1} - \lambda_{2r-1}).\]

(3.90)
\textbf{Proof} Let $w_{11}$ and $w_{22}$ denote the diagonal entries of $V_{r,c}^{off}$ in (3.73). Note that $w_{11}$ and $w_{22}$ are almost identical to the diagonal elements $v_{11}$ and $v_{22}$ of its counterpart $V_{r,c}^{diag}$, except the sums of $\lambda_i$'s appear in different pairs:

\[
w_{11} = (\lambda_{2r-1} + \lambda_{2c}) + \langle \Phi', \widehat{E} \rangle \\
= \frac{(a_r^2 - a_c^2)(\lambda_{2r-1} + \lambda_{2c})}{a_r^2 - a_c^2} + \frac{-a_r^2(\lambda_{2r-1} + \lambda_{2r}) + a_c^2(\lambda_{2c-1} + \lambda_{2c})}{a_r^2 - a_c^2} \\
= \frac{(\lambda_{2c-1} - \lambda_{2r-1})a_r^2 + (\lambda_{2c} - \lambda_{2r})a_c^2}{a_r^2 - a_c^2}, \quad (3.91)
\]

and likewise

\[
w_{22} = (\lambda_{2r} + \lambda_{2c-1}) + \langle \Phi', \widehat{E} \rangle = \frac{(\lambda_{2c-1} - \lambda_{2r-1})a_r^2 + (\lambda_{2c} - \lambda_{2r})a_c^2}{a_r^2 - a_c^2}. \quad (3.92)
\]

Now we calculate the product $w_{11}w_{22}$ of diagonal entries in $V_{r,c}^{off}$ similarly to the product (3.88),

\[
w_{11}w_{22} = \left[(\lambda_{2r-1} + \lambda_{2c}) + \langle \Phi', \widehat{E} \rangle \right]\left[(\lambda_{2r} + \lambda_{2c-1}) + \langle \Phi', \widehat{E} \rangle \right]
\[
= \left(\frac{\lambda_{2c-1} - \lambda_{2r-1}a_r^2 + (\lambda_{2c} - \lambda_{2r})a_c^2}{a_r^2 - a_c^2}\right)\left(\frac{\lambda_{2c-1} - \lambda_{2r-1})a_r^2 + (\lambda_{2c} - \lambda_{2r})a_c^2}{a_r^2 - a_c^2}\right)
\[
= \frac{(\lambda_{2c-1} - \lambda_{2r-1})^2 + (\lambda_{2c} - \lambda_{2r})^2}{(a_r^2 - a_c^2)^2}a_d^2a_e^2 + (\lambda_{2c-1} - \lambda_{2r-1})a_c^2 + (\lambda_{2c} - \lambda_{2r})(a_e^2 + a_d^2)}{a_r^2 - a_c^2} \quad (3.93)
\]

The product of anti-diagonal elements is the exact same as for $V_{r,c}^{diag}$, namely

\[
[- \langle \Phi', \widehat{O} \rangle]^2 = \left[\langle \Phi', \widehat{O} \rangle \right]^2 \quad \text{computed in (3.89); however, it is more convenient in this case to group the terms as follows}
\]

\[
[- \langle \Phi', \widehat{O} \rangle]^2 = \left(-a_e a_c (\lambda_{2r-1} + \lambda_{2c}) + a_r a_c (\lambda_{2c-1} + \lambda_{2c}) \right)^2
\[
= \left(\frac{(\lambda_{2c} - \lambda_{2r})a_r a_e + (\lambda_{2c-1} - \lambda_{2r-1})a_r a_c}{a_r^2 - a_c^2}\right)^2
\[
= \frac{(\lambda_{2c} - \lambda_{2r})^2 + (\lambda_{2c-1} - \lambda_{2r-1})^2 + 2(\lambda_{2c} - \lambda_{2r})(\lambda_{2c-1} - \lambda_{2r-1})}{(a_r^2 - a_c^2)^2}a_d^2a_e^2. \quad (3.94)
\]

By arranging terms in (3.94), the term $[(\lambda_{2c} - \lambda_{2r})^2 + (\lambda_{2c-1} - \lambda_{2r-1})^2]a_d^2a_e^2$ is can-
celled in (3.93), and therefore we can simplify the determinant of $V_{r,c}^{\text{off}}$,  

$$
\det \left( V_{r,c}^{\text{off}} \right) = w_{11}w_{22} - \left[ -s \right]^2 \\
= \frac{(\lambda_{2c} - \lambda_{2r})(\lambda_{2c-1} - \lambda_{2r-1})(a_r^2 + a_c^2) - 2(\lambda_{2c} - \lambda_{2r})(\lambda_{2c-1} - \lambda_{2r-1})a_r^2a_c^2}{(a_r^2 - a_c^2)^2} \\
= \left( \frac{a_r^4 - 2a_r^2a_c^2 + a_c^4}{(a_r^2 - a_c^2)^2} \right)(\lambda_{2c} - \lambda_{2r})(\lambda_{2c-1} - \lambda_{2r-1}) \\
= (\lambda_{2c} - \lambda_{2r})(\lambda_{2c-1} - \lambda_{2r-1}).$$

Theorems 3.4.5 and 3.4.6 thus imply conditions for the determinants of $V_{r,c}^{\text{diag}}$ and $V_{r,c}^{\text{off}}$ to both be positive, a necessary condition for $V_{r,c}^{\text{diag}}$ and $V_{r,c}^{\text{off}}$ to be either positive or negative definite.

**Corollary 3.4.7** Given any pair $(r, c)$ such that $1 \leq r < c \leq n$, the determinants of the blocks $V_{r,c}^{\text{diag}}$ and $V_{r,c}^{\text{off}}$ are positive if and only if the corresponding entries $\lambda_{2r-1}$, $\lambda_{2r}$, $\lambda_{2c-1}$, and $\lambda_{2c}$ of the inertia matrix $J$ satisfy one of the following two sets of inequalities:

$$
\lambda_{2r-1} < \lambda_{2c-1}, \lambda_{2c} \\
\lambda_{2r} < \lambda_{2c-1}, \lambda_{2c} \tag{3.95}
$$

or

$$
\lambda_{2r-1} > \lambda_{2c-1}, \lambda_{2c} \\
\lambda_{2r} > \lambda_{2c-1}, \lambda_{2c}. \tag{3.96}
$$

**Proof** In order for the determinants of $V_{r,c}^{\text{diag}}$ and $V_{r,c}^{\text{off}}$ to be positive, it is clear from Theorems 3.4.5 and 3.4.6 that the two inequalities below must hold:

$$
(\lambda_{2c} - \lambda_{2r-1})(\lambda_{2c-1} - \lambda_{2r}) > 0 \\
(\lambda_{2c} - \lambda_{2r})(\lambda_{2c-1} - \lambda_{2r-1}) > 0. \tag{3.97}
$$

Assume $\lambda_{2c} > \lambda_{2r-1}$ in which case we must have $\lambda_{2c-1} > \lambda_{2r}$. Next, if $\lambda_{2r} > \lambda_{2c}$, the second inequality is satisfied only if $\lambda_{2r-1} > \lambda_{2c-1}$. We have arrived at a
contradiction since
\[ \lambda_{2c} > \lambda_{2r-1} > \lambda_{2c-1} > \lambda_2 > \lambda_{2c}. \]
Thus if \( \lambda_{2c} > \lambda_{2r-1} \), it follows that \( \lambda_{2c} > \lambda_{2r} \) and \( \lambda_{2c-1} > \lambda_{2r-1} \), \( \lambda_{2r} \).

Conversely if \( \lambda_{2c} < \lambda_{2r-1} \), then it immediately follows that \( \lambda_{2c-1} < \lambda_{2r} \). Now assume \( \lambda_{2c} > \lambda_{2r} \) and thus \( \lambda_{2c-1} > \lambda_{2r-1} \). These conditions lead to yet another contradiction since
\[ \lambda_{2c} < \lambda_{2r-1} < \lambda_{2c-1} < \lambda_{2r} < \lambda_{2c}. \]

Therefore if \( \lambda_{2c} < \lambda_{2r-1} \), inequalities (3.97) are satisfied if and only if the ordering of the \( \lambda_i \)'s satisfy the condition in (3.96). Altogether we have now proved the determinants of \( V_{r,c}^{\text{diag}} \) and \( V_{r,c}^{\text{off}} \) are both positive if and only if exactly one of the conditions (3.95) or (3.96) are satisfied.

Together Corollaries 3.4.3 and 3.4.7 provide necessary and sufficient conditions on the \( \lambda_i \)'s such that each consecutive pair \( V_{r,c}^{\text{diag}} \) and \( V_{r,c}^{\text{off}} \) are both positive or negative definite. Recall that there are \((n^2 - n)\) blocks \( M_{r,c} \) for which we get a corresponding pair of \( 2 \times 2 \) blocks \( V_{r,c}^{\text{diag}} \) and \( V_{r,c}^{\text{off}} \) along the diagonal of the second variation matrix \( d^2 H_\phi(\Omega_x) \). Thus in order for all of the \((n^2 - n)\) pairs of \( 2 \times 2 \) blocks to have positive traces, the \( \lambda_i \)'s must simultaneously satisfy equation (3.79) for all \( n^2 - n \) pairs \((r, c)\) such that \( 1 \leq r < c \leq n \). The determinants will in this case be positive if and only if the inequalities (3.96) hold for each of the \((n^2 - n)\) pairs of \( r \) and \( c \).

**Theorem 3.4.8** The \( 2 \times 2 \) blocks are all positive definite if and only if the \( \lambda_i \)'s in
the inertia matrix $J$ satisfy the ordering

$$
\lambda_1 > \lambda_3, \lambda_4, \ldots, \lambda_{2n} \\
\lambda_2 > \lambda_3, \lambda_4, \ldots, \lambda_{2n} \\
\lambda_3 > \lambda_5, \lambda_6, \ldots, \lambda_{2n} \\
\lambda_4 > \lambda_5, \lambda_6, \ldots, \lambda_{2n} \\
\vdots \\
\lambda_{2n-3} > \lambda_{2n-1}, \lambda_{2n} \\
\lambda_{2n-2} > \lambda_{2n-1}, \lambda_{2n}.
$$

Conversely the traces of all $2 \times 2$ blocks are negative if and only if all corresponding inequalities (3.80) are satisfied, in which case the determinants are positive if and only if the conditions in (3.95) hold for each of the $(n^2 - n)$ pairs of $r$ and $c$.

**Theorem 3.4.9** The $2 \times 2$ blocks are all negative definite if and only if the $\lambda_i$'s in the inertial matrix $J$ satisfy the following ordering,

$$
\lambda_1 < \lambda_3, \lambda_4, \ldots, \lambda_{2n} \\
\lambda_2 < \lambda_3, \lambda_4, \ldots, \lambda_{2n} \\
\lambda_3 < \lambda_5, \lambda_6, \ldots, \lambda_{2n} \\
\lambda_4 < \lambda_5, \lambda_6, \ldots, \lambda_{2n} \\
\vdots \\
\lambda_{2n-3} < \lambda_{2n-1}, \lambda_{2n} \\
\lambda_{2n-2} < \lambda_{2n-1}, \lambda_{2n}.
$$

### 3.4.2 Definiteness of the $n \times n$ Block

Now we turn our attention to the remaining $n \times n$ block $T_{n \times n} = V_{n \times n} + S_{n \times n}$ of $d^2H_\phi(\Omega_e)$ given in (3.68). Recall the block $T_{n \times n}$ gives the value of the second variation $d^2H_\phi(\Omega_e)$ at a pair of matrices $(Y, Z)$ in the Cartan subalgebra $\mathfrak{h}$. Finding the eigenvalues of $T_{n \times n}$ is quite overwhelming. Instead, we will prove the definiteness
of $T_{n \times n}$ by first noting if $Hess(\phi)$ is definite, then block $S_{n \times n} = d\tilde{F}_{\Omega_e}^t Hess(\phi) d\tilde{F}_{\Omega_e}$ is definite as well. In general, the matrix $V_{n \times n}$ is indefinite. However, we will prove it is possible to find conditions on the Hessian of $\phi$ at $F(\Omega_e)$ such that the definiteness of the matrix $S_{n \times n}$ is enough to guarantee the definiteness of the sum $T_{n \times n} = V_{n \times n} + S_{n \times n}$. It is important to note that the conditions on $Hess(\phi)$ will not contradict any of the previous restrictions on $\phi$ necessary and sufficient for $dH_\phi(\Omega_e) = 0$ (see Theorem 3.2.4), and the definiteness of $T_{n \times n}$ will not impose any further conditions on the $\lambda_i$’s.

We now focus on finding conditions on the function $\phi$ that will ensure

$$T_{n \times n}(Y, Y) = Y^t T_{n \times n} Y > 0 \quad \text{or} \quad T_{n \times n}(Y, Y) = Y^t T_{n \times n} Y < 0 \quad (3.100)$$

for all matrices $Y = \sum_{i=1}^n (c_i Y^i_d)$ (with $c_i \in \mathbb{R}$ for $1 \leq i \leq n$) corresponding to matrices from the Cartan subalgebra $\mathfrak{h}$ defined in (2.4). First, note that it is sufficient to check the conditions in (3.100) for all matrices $Y \in \mathfrak{h}$ with $\|Y\| = \sqrt{<Y, Y>} = 1$ (where $<,>$ is the trace inner product in (3.6)), since for any scalar $c$, $T_{n \times n}(cY, cY) = c^2 T_{n \times n}(Y, Y)$. Next, we note that if $M$ is a positive (negative) definite, and $Q$ is any invertible matrix, then it must follow that $Q^* M Q$ is positive (negative) definite as well.

Recall the Jacobian $d\tilde{F}_{\Omega_e}$ defined in Theorem 3.2.4. Since the generic equilibrium $\Omega_e$ lies on a non-degenerate adjoint orbit, it is a regular point of $\tilde{F}$. Thus the Jacobian $d\tilde{F}$ is invertible at $\Omega_e$, and the matrix $(d\tilde{F}_{\Omega_e}^{-1})^t T_{n \times n} d\tilde{F}_{\Omega_e}^{-1}$ exists. Next, we prove definiteness of $(d\tilde{F}_{\Omega_e}^{-1})^t T_{n \times n} d\tilde{F}_{\Omega_e}^{-1}$, and observe that definiteness of $T_{n \times n}$ will then follow.

In Theorem 3.2.4 we found conditions on the first partial derivatives of $\phi$ that make $\Omega_e$ a critical point of $\phi_F$: when the column vector of first partial derivatives $\Phi' = -\left(d\tilde{F}_{\Omega_e}^t\right)^{-1} AA$. 
Lemma 3.4.10  For any $c \in \mathbb{R}\{0\}$, the function
\[
\phi : \mathbb{R}^n \to \mathbb{R} : x = (x_1, \ldots, x_n) \mapsto x \cdot \Phi' + c \left( \sum_{i=1}^{n} (x_i - C_i(\Omega_e))^2 \right) \tag{3.101}
\]
will have a critical point at $F(\Omega_e)$ and Hessian
\[
\text{Hess}(\phi) = c (I_{n \times n}) . \tag{3.102}
\]

Proof  The proof of the lemma follows immediately from the construction of the function. Clearly the Hessian of $\phi$ is $c(I_{n \times n})$ since
\[
\phi_{ij}(F(\Omega_e)) = \begin{cases} 
0 & i \neq j \\
c & i = j 
\end{cases} .
\]

Next, taking the $k^{th}$ partial derivative of $\phi$ in (3.101), we obtain
\[
\phi_k(x) = \phi_k(F(\Omega_e)) + 2c (x_k - C_k(\Omega_e)) \quad \text{for all } 1 \leq k \leq n.
\]

Note that in the expression above, the term $\phi_k(F(\Omega_e))$ is the constant number determined by the formula $\Phi' = - (dF_{\Omega_e}^t)^{-1} A \Lambda$ found in Theorem 3.2.4. Thus, we have constructed our function $\phi$ in (3.101) such that at $x = F(\Omega_e) = (C_1(\Omega_e), \ldots, C_n(\Omega_e))$, it directly follows that the $k^{th}$ partial derivative of $\phi$ is indeed the $k^{th}$ entry of the column vector $\Phi'$ in (3.16).

The matrix $T_{n \times n}$ is equal to the sum of matrices $S_{n \times n}$ and $V_{n \times n}$ found in Theorem 3.3.8 and Corollary 3.63 respectively. For convenience we set $Q = dF_{\Omega_e}^{-1}$ and express $S_{n \times n} = (Q^{-1})^t \text{Hess}(\phi) Q^{-1}$. Given the $\phi$ in (3.101), with $\text{Hess}(\phi) = c (I_{n \times n})$, it immediately follows that $S_{n \times n}$ is positive definite if $c > 0$ and negative definite if $c < 0$. The matrix $V_{n \times n}$ is in general indefinite; however the following theorem proves it is possible to choose a coefficient $c$ in (3.101) such that the definiteness of
the matrix $S_{n \times n}$ will dominate the matrix $V_{n \times n}$. In the statement and proof of the theorem, we use the notation $M(Y, Y)$ to denote $Y'MY$ and recall $\|Y\| = \sqrt{\langle Y, Y \rangle}$ denotes the norm of the matrix $Y$ with respect to the trace form introduced in (3.6).

**Theorem 3.4.11** If $c > \max_{\|Y\|=1} \left| (Q^t V_{n \times n} Q) (Y, Y) \right|$, the matrix $T_{n \times n}$ is positive definite.

**Proof** The norm of $Q^t V_{n \times n} Q$ is bounded; thus there exists a positive, real number $M$ such that

$$\max_{\|Y\|=1} \left| (Q^t V_{n \times n} Q) (Y, Y) \right| \leq M \quad \text{for all } \|Y\| = 1.$$  

(3.103)

Since $S_{n \times n} = (Q^{-1})^t Hess(\phi) Q^{-1}$, it clearly follows that

$$(Q^t S_{n \times n} Q) (Y, Y) = Hess(\phi)(Y, Y) = c\|Y\|$$

Now we find a value for $c$ such that for all $\|Y\| = 1$ we have $$(Q^t T_{n \times n} Q) (Y, Y) > 0:$$

$$\min_{\|Y\|=1} \left( Q^t T_{n \times n} Q \right) (Y, Y) = \min_{\|Y\|=1} \left\{ (Q^t V_{n \times n} Q) (Y, Y) + (Q^t S_{n \times n} Q) (Y, Y) \right\}$$

$$\geq \left( \min_{\|Y\|=1} \left( Q^t V_{n \times n} Q \right) (Y, Y) \right) + \left( \min_{\|Y\|=1} \left( Q^t S_{n \times n} Q \right) (Y, Y) \right)$$

$$\geq -\max_{\|Y\|=1} \left| (Q^t V_{n \times n} Q) (Y, Y) \right| + \left( \min_{\|Y\|=1} \left( Hess(\phi) \right) (Y, Y) \right)$$

$$= -M + c.$$

Thus, if $c > \max_{\|Y\|=1} \left| (Q^t V_{n \times n} Q) (Y, Y) \right| = M$, then $\min_{\|Y\|=1} \left( Q^t T_{n \times n} Q \right) (Y, Y) > 0$, and $T_{n \times n}$ is positive definite. \[\blacksquare\]

**Corollary 3.4.12** A given generic equilibrium $\Omega_e$ of the form (2.19) is Lyapunov stable if the $\lambda_i$’s satisfy the ordering in Theorem 3.4.8.
Similarly if we wish to show that $T_{n \times n}$ is negative definite, notice that for any $c < -\max_{\|Y\|=1} \left| \left( Q^t V_{n \times n} Q \right) (Y, Y) \right|$ it follows that $\max_{\|Y\|=1} T_{n \times n}(x) < 0$. Therefore if the $\lambda_i$’s satisfy the ordering in Theorem 3.4.9, constructing a $\phi$ in Lemma 3.4.10 using this $c$ yields a Hamiltonian $H_\phi$ which has a critical point at $\Omega_e$, whose second variation is negative definite at $\Omega_e$.

**Corollary 3.4.13** *A given generic equilibrium $\Omega_e$ of the form (2.19) is Lyapunov stable if the $\lambda_i$’s satisfy the ordering in Theorem 3.4.9.*
Euler’s derivation of the equations of motion for the three dimensional rigid body can be traced back to his work in 1750 of *Discovery of a new principle of mechanics* and the eventual publication of his solution to the three dimensional rigid body problem in 1760 in his *Theoria motus corporum solidorum seu rigidorum* [9]. Euler’s classical equations for the three dimensional rigid body are

\[ I_1 \dot{\omega}_1 = (\lambda_2 - \lambda_3)\omega_2\omega_3 \]
\[ I_2 \dot{\omega}_2 = (\lambda_3 - \lambda_1)\omega_1\omega_3 \]
\[ I_3 \dot{\omega}_3 = (\lambda_1 - \lambda_2)\omega_1\omega_2. \]

Essential to Euler’s work on the rigid body is his understanding of the principle moments of inertia and the use the of a moving frame of reference fixed in the body [9].

Based on the his discovery, Euler observes that the only way a rigid body will rotate along fixed axis is when the axis is one of the principal axes, thus finding all stable solutions. Euler continues to find solutions corresponding to specific cases, but the general solution to the problem evades Euler as the equations are quite complicated (involving elliptic functions). Euler wisely predicts that the solution to the problem should be approached by means other than analytic tools of mechanics [8].

Euler’s legacy continues to grow with the work of his student Lagrange. Lagrange continued investigating the mechanics of the rigid body in the hope of “reducing the theory of this science [mechanics], and the art of solving the problems concerned
with it, to general formulae, the simple development of which gives all the equations necessary to the solution of each problem.” [29]. In his major work *Mécanique Analytique* published in 1788, Lagrange formalizes the work of D’Almbert and Euler and uncovers the general principle of least action. Using purely an analytic approach, Lagrange derives the universal Euler-Lagrange equation. Lagrange’s approach is applied to the rigid body, and Euler’s equation is restated conveniently as a cross product of the momentum and velocity:

\[
\frac{d}{dt} (\mathbb{I} \omega) = (\mathbb{I} \omega) \times \omega
\]

Generally credited as being the father of geometric mechanics, Poinsot derives an elegant geometrical description of the motion of the rigid body, without requiring explicit solutions to the equations, in his 1834 work *Théorie nouvelle de la rotation des corps*. The fact that Poinsot’s method is still discussed in most classical mechanics texts is a tribute to the beauty and simplicity of his solution. Using the conservation of energy and angular momentum, Poinsot observes the motion of the rigid body can be illustrated by an inertia ellipsoid (corresponding to the energy) which rolls on an invariant plane whose normal vector is constant (corresponding to the fixed angular momentum in space). The motion of the three dimensional rigid body is thus equivalent to the study of the curves traced in the invariant plane by the point of contact on the ellipsoid (see figure A). Poinsot is then able to show rotations along either the longest or shortest axes of the inertia ellipsoid are stable equilibria, while rotations along the intermediate axis are unstable equilibria.

Shortly after Euler and Lagrange build a mathematical foundation for mathematics of Newtonian mechanics, Hamilton develops the mathematics of optics in his series of papers *Theory of Systems of Rays* beginning as an undergraduate in 1826. He formalizes his notation of a characteristic function, or the appropriately named Hamiltonian, and its application in the study of optics in one of his major works, *On*
**Figure A.1**: An illustration of Poinsot’s construction

*a General Method in Dynamics* in 1834, the same year Poinsot derives his geometric methods to study the dynamics of the rigid body. Motivated by Hamilton’s work, Jacobi imports Hamilton’s theories on the dynamics of optics into mechanics and creates the Hamilton-Jacobi formalism in mechanics. Merging his work on elliptic functions, Jacobi is famous for discovering in 1839 that the geodesic motion on an ellipsoid is solvable by hyperelliptic functions. Liouville next formalizes the concept of integrability introduced by Jacobi. Shortly thereafter in 1889, Kowalevski proves that the rigid body is a completely integrable system, which has since led to many exciting connections to algebraic geometry and hyperelliptic functions.

After successful developments in the dynamics of Euler’s three dimensional rigid body, the problem is considered in higher dimensional space. Namely what are the corresponding equations for the rigid body which rotates freely about a fixed point in $n$ dimensional space. The origins of the $n$ dimensional generalization can be traced back to Cayley in 1846, and equations for the $n$ dimensional generalizations of Euler’s classical rigid body are discovered in 1875 by Frahm. As early as 1891, the integrability of the 4 dimensional rigid body is proved by Schottky [6].

Though techniques for integrating many different ordinary differential equations have made major strides by the end of the 19th century, the various methods and
problems seemed internally disconnected. Lie amazingly derives a majority of these methods by applying the general theory of Lie groups. Lie continues to classify all such ordinary differential equations in terms of the symmetry groups they admit. In the process, Lie entirely describes the set of equations whose solutions can be obtained by applying properties from the underlying groups to reduce the order of the systems [15]. Though Lie develops his ideas apart from mechanics, the gap between mechanics and geometry is closing.

At the turn of the century a major breakthrough occurs: Poincare discovers that the mechanics of rigid bodies and fluids could be expressed in the setting of Lie algebras, and thus generalizes the work of Euler and Lagrange into the context of Lie’s work with the publication of his Euler-Poincare equation in 1901 [23]. Poincare derives the equations of such systems as the rigid body in the dual space of the corresponding Lie algebras of the symmetry groups. The implications of Poincare’s generalization become increasingly apparent as the the concepts of Lie are developed in the beginning of the 20th century.

In 1966 Arnold revolutionizes the applications of modern geometry which followed from Lie’s discoveries. Arnold publishes his beautiful paper Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l’hydrodynamique des fluides parfaits on the two hundredth anniversary of the original publication of Euler’s equation for the motion of three dimensional rigid body [2]. In this paper, Arnold exploits the role of Lie groups in the dynamics of systems defined on Lie groups, such as the rigid body. Arnold defines the equations for generalized n dimensional rigid body in the dual Lie algebra using the more developed concepts of Lie,

\[ \dot{M} = ad^*_{\exp(\Omega t)} M, \]

for \( M \) and \( \Omega \) the angular momentum and velocity with respect to the moving frame of the body. Arnold also lays the groundwork for the energy-Casimir method by introducing a modified Lyapunov method for studying the stability of equilibria in
reduced systems on dual Lie algebras.

In the 1970’s, Marsden and Weinstein extend Arnold’s work to the study of Hamiltonian systems with symmetry, whose configuration space is not necessarily a Lie group. In 1974, they publish a seminal paper [22] in which they describe symplectic reduction for systems with symmetry. This method has become a cornerstone for modern geometric mechanics.
We now include the details of the calculation of the entries of the matrix \( d\tilde{F}_{\Omega_e} ^t \, Hess(\phi) \, d\tilde{F}_{\Omega_e} \), where the Hessian of \( \phi \) at \( \tilde{F}(\Omega_e) \) and the Jacobian \( d\tilde{F}_{\Omega_e} ^t \) are given by

\[
Hess(\phi) = \begin{pmatrix}
\phi_{11} & \phi_{12} & \phi_{13} & \ldots & \phi_{1n} \\
\phi_{12} & \phi_{22} & \phi_{23} & \ldots & \phi_{2n} \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\phi_{1,n-1} & \phi_{2,n-1} & \phi_{3,n-1} & \ldots & \phi_{n-1,n} \\
\phi_{1n} & \phi_{2n} & \phi_{3n} & \ldots & \phi_{nn}
\end{pmatrix}
\]

and

\[
d\tilde{F}_{\Omega_e} ^t = \begin{pmatrix}
a_1 & a_1^3 & \ldots & a_2 a_3 \ldots a_n \\
a_2 & a_2^3 & \ldots & a_1 a_3 \ldots a_n \\
\vdots & \ddots & \ddots & \ddots \\
a_n & a_n^3 & \ldots & a_1 a_2 \ldots a_{n-1}
\end{pmatrix}
\]

respectively. The \( ij^{th} \) entry, \( \left( d\tilde{F}_{\Omega_e} ^t \, Hess(\phi) \, d\tilde{F}_{\Omega_e} \right)_{ij} \), is computed by multiplying \( Hess(\phi) \) on the left by the \( i^{th} \) row of \( d\tilde{F}_{\Omega_e} ^t \), and on the right by the \( j^{th} \) column of \( d\tilde{F}_{\Omega_e} ^t \).
\[
\begin{align*}
(dF^{t}_{\Omega} Hess(\phi) dF_{\Omega})_{ij} &= \left(\begin{array}{cccc}
a_i & a_i^3 & \cdots & a_i^{2n-3} \\
a_i^3 & \phi_{11} & \phi_{12} & \cdots & \phi_{1n} \\
\vdots & \vdots & \ddots & \vdots \\
a_i^{2n-3} & \phi_{1,n-1} & \phi_{2,n-1} & \cdots & \phi_{n-1,n} \\
\frac{b}{a_i} & \phi_{1n} & \phi_{2n} & \cdots & \phi_{nn}
\end{array}\right) \times \left(\begin{array}{c}
a_j \\
a_j^3 \\
\vdots \\
a_j^{2n-3} \\
\frac{b}{a_j}
\end{array}\right) \\
&= \left(\begin{array}{c}
\sum_{k=1}^{n-1} (a_i^{2k-1} \phi_{1k}) + \frac{b}{a_i} a_j \phi_{1n} \\
\sum_{k=1}^{2} (a_i^{2k-1} \phi_{k2}) + \frac{b}{a_i} \phi_{2n} \\
\vdots \\
\sum_{k=1}^{n-1} (a_i^{2k-1} \phi_{kn}) + \frac{b}{a_i} \phi_{nn}
\end{array}\right) \\
&= \sum_{k=1}^{n-1} \left( a_j a_i^{2k-1} \phi_{1k} \right) + \frac{b}{a_i} a_j \phi_{1n} \\
&\quad + \sum_{k=1}^{n-1} \left( a_j^3 a_i^{2k-1} \phi_{k2} \right) + \frac{b}{a_i} a_j^3 \phi_{2n} + \cdots \\
&\quad + \sum_{k=1}^{n-1} \left( \frac{b}{a_j} a_i^{2k-1} \phi_{kn} \right) + \frac{b^2}{a_i a_j} \phi_{nn} \\
&= \sum_{l=1}^{n-1} a_j^{2l-1} \sum_{k=1}^{n-1} (a_i^{2k-1} \phi_{lk}) + \sum_{l=1}^{n-1} \left( a_j^{2l-1} \frac{b}{a_i} \phi_{ln} \right) \\
&\quad + \sum_{k=1}^{n-1} \left( \frac{b}{a_j} a_i^{2k-1} \phi_{kn} \right) + \frac{b^2}{a_i a_j} \phi_{nn}
\end{align*}
\]
REFERENCES


