THE UNIVERSITY OF TEXAS AT AUSTIN

Riemann Ellipsoids: Hamiltonian Formulation and Stability Analysis

by

Santiago J. Benavides

In partial fulfillment of the requirements for graduation with the Dean's Scholars Honors Degree in Physics and Mathematics

under the supervision of
Dr. P. J. Morrison
and
Dr. I. M. Gamba
in the
Departments of Physics and Mathematics
College of Natural Sciences

May 2015

In partial fulfillment of the requirements for graduation with the
Dean's Scholars Honors Degree in Physics and Mathematics

Santiago José Benavides Dean's Scholars Honors Program	Date
Prof. Philip J. Morrison Supervising Professor Department of Physics	Date
Prof. Irene M. Gamba Co-Supervising Professor Department of Mathematics	Date
Prof. Greg Sitz Honors Advisor in Physics	Date

I grant the Dean's Scholars Program permission to post
a copy of my thesis on the University of Texas Digital Repository.
For more information, visit http://repositories.lib.utexas.edu/about.

Santiago José Benavides — Date

Signature

Prof. Philip J. Morrison — Date

Signature

Prof. Irene M. Gamba — Date

Signature

Declaration of Authorship

I, SANTIAGO J. BENAVIDES, declare that this thesis titled, 'RIEMANN ELLIPSOIDS: HAMILTONIAN FORMULATION AND STABILITY ANALYSIS' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Date: _____

NU PASSARIELLO SPIERZO

Nu passariello spierzo e abbandunato
'ncopp'a na casa janca se pusaie.
Chi nc'era? Nu guaglione allora nato!
Tante d' 'e strille, 'o passero tremmaie.
E senza sciato, muorto d' 'a paura,
lassaie sta casa janca e 'a criatura,

Jette a pusarze 'mmiez'a na furesta
tutta friscura e tutt'erba addurosa.
Vedenno primma n'ommo e po' na vesta
'o passero capette quacche cosa...
Ma quanno lle dicettero: Vattenne!
se ne scappaie, lassanno llà doie penne.

Cchiù tarde, se pusaie 'ncopp'a na cchiesa
addò nce steva 'a ggente 'a cchiù sfarzosa.
Spusavano nu conte e na marchesa;
chiagnevano (pecché?) o sposo e a sposa,
'O passero penzaie: - Vi' ch'allegria! -
e ghiette a cercà n'ata cumpagnia.

All'urderm'ora, po', na capannella
zitta e sulanga 'o passero truvaie:
nce steva, morta già, na vicchiarella
ch'era campata senz'ammore e guaie.
E 'ncopp'a sta capanna accujetata
'o passero passaie tutt' 'a nuttata.

Roberto Bracco

THE UNIVERSITY OF TEXAS AT AUSTIN

Abstract

Dr. P. J. Morrison
Dr. I. M. Gamba
Departments of Physics and Mathematics
College of Natural Sciences

Bachelor of Science in Physics and Mathematics (Option: Honors)

by Santiago J. Benavides

The equilibria and stability of self-gravitating liquid masses has been studied and debated for more than a century by great physicists and mathematicians such as Newton, Maclaurin, Jacobi, Poincaré, Dirichlet, Riemann, and Chandrasekhar, and is still drawing interest from researchers today. Here I present an original approach to formulating the problem in the context of Hamiltonian theory, namely by applying moments of the position and velocity to the constrained Poisson bracket for a fluid. I then study the stability of a certain family of equilibrium ellipsoids with internal flow that depends linearly on the spatial coordinates (Riemann ellipsoids) using this constrained Hamiltonian formulation of the problem. This formulation allows us to use robust stability analysis methods, as well as study the dynamics in a straightforward way. The spectral stability results agree qualitatively with that of Chandrasekhar's, but the parameter value is slightly off, and the nonlinear stability analysis results do not give a definite answer due to the nature of the bifurcation (steady-state). It is still possible to use the Rayleigh-Ritz method to determine whether our system is nonlinearly unstable, but due to time constraints this was not done.

Acknowledgements

I would like to thank Dr. Philip J. Morrison (my advisor) for his incredible generosity and incomparable enthusiasm towards the subject and being a teacher. Without the many hours of group meetings, individual lessons, and discussions over the years that we've had, I would not be half the student I am now and certainly I would not be as curious about the many things that interest me at the moment. I would also like to thank Dr. Norm Lebovitz of the University of Chicago, with whom (along with Dr. Morrison) I began this project in the summer of 2014 at Woods Hole Oceanographic Institution's "Geophysical Fluid Dynamics" program in Woods Hole, MA. Finally, I would like to thank Dr. Irene Gamba for her generosity and being of much help during the process of writing this thesis.

Another huge influence was Dr. Alexandros Alexakis, who was one of my first true role models as a teacher, scientist, and person. He took me in as a naïve sophomore who knew nothing about fluid dynamics, or what research was about, and had the patience and generosity to take time out of his day to teach me what he knew. I learned so much in the three life-changing months that I spent in Paris, and it certainly shaped who I am today, academically and personally.

I am very grateful for Dr. Harry Swinney and his group, Dr. Likun Zhang, Dr. Bruce Rodenborn and Dr. Michael Allshouse, for letting me work in their lab and bother them with questions and concerns. I learned so much from the group meetings and from the presentations I witnessed (and from the ones I made, as well!). Working in this lab made me realize how important good communication is in science. It also made me appreciate the essential relationship that theorists and experimentalists have, and the potential to do great science together.

Of course, a great community of friends is necessary in order for sanity to be maintained and life to be enjoyed. Therefore, I want to thank Danny, Jay, Jacob, Jules, Sean, Gabe, Eric, Manasvi, Akarsh, George, Behzat, Siva, and Rustam (to name a few) for being great friends, role models, and companions during my studies at UT. I couldn't've done it without the late night homework sessions, the Whizkey talks, the HULS talks, the dinners and many other great times.

And finally, it goes without saying that my undergraduate experience was not all research-based, but also had a lot to do with the (mostly) amazing teachers I had. I would like to thank the teachers who have inspired me to learn more but also how important it is to teach – Dr. Greg Sitz, Dr. Can Kilic, Dr. Sonia Paban, Dr. David Laude, Dr. Phil Morrison, Dr. Dan Freed, Dr. Sean Keel, Dr. Harry Swinney, and more.

Contents

List of Figures

Lo dedico a mis padres, Ida y Fernando. Los amo!

Chapter 1

Introduction

1.1 Historical Context

I will start with a question posed by E. T. Whittaker in a lecture he gave titled "Spin in the Universe," and quoted by Chandrasekhar in his book [4]:

> Rotation is a universal phenomenon; the earth and all the other members of the solar system rotate on their axes, the satellites revolve round the planets, the planets revolve round the Sun, and the Sun himself is a member of the galaxy or Milky Way system which revolves in a very remarkable way. How did all these rotatory motions come into being? What secures their permanence or brings about their modification? And what part do they play in the system of the world?

As Whittaker implies, rotation is seen to exist unanimously throughout our universe; and it is not without consequences! Even focusing on our local neighborhood, we see that the solar system and our experience in it as we know it would be entirely different without this phenomenon. Off the bat, one can see that we would have no concept of days or years, and the climate and weather would be drastically different, as well. Even more importantly there would be no accretion disk to form our solar system, so forget about having planets! All of these details, although interesting, are rather obvious or too hypothetical to be relevant. However, there are more subtle results of rotation that are worth asking. One particular example that has been asked for more than a century is the question of the *shape* of a rotating object. How does rotation affect the shape of an object whose surface is not rigid (a "free surface")? For example, a rotating drop of water or even a crude approximation of the Earth or Sun.

This is what Newton first posed in his monumental book *Principia* (1687) – he asked about the shape of the Earth, given that it was spinning on its axis. He decided to tackle this question by first assuming that the Earth was a *homogeneous self-gravitating fluid*[1], and then inquired about its shape. With this he predicted the eccentricity ϵ of the Earth, defined to be

$$\epsilon = \frac{\text{equatorial radius} - \text{polar radius}}{\text{mean radius}(R)},$$

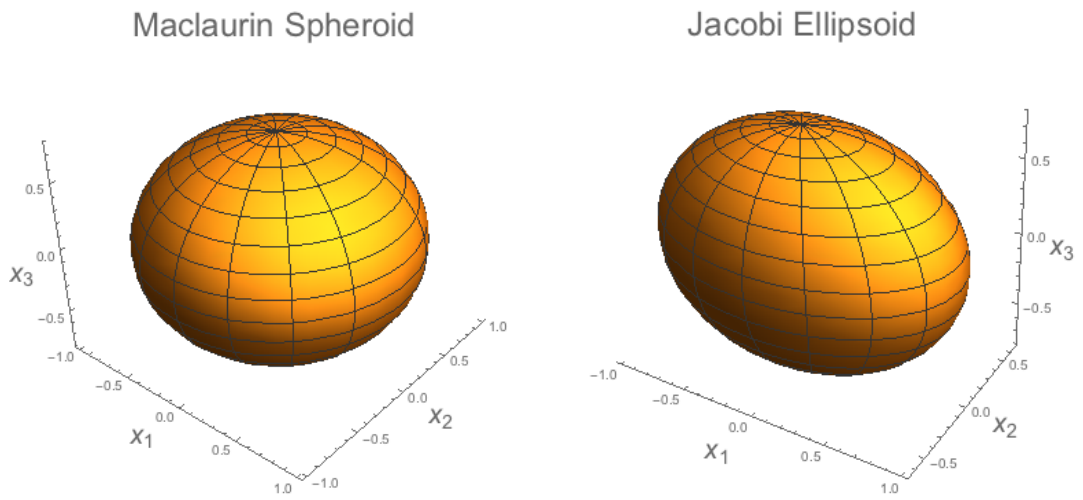
assuming that the equilibrium shape would be that of an oblate spheroid (flattened along axis of rotation, rotationally symmetric), and got $\epsilon \sim 1/230$ which is surprisingly close to the current estimate of $\epsilon \sim 1/294$ [4]. However, this shape was merely an educated guess as to what an equilibrium shape of a spinning, self-gravitating object would look like (in fact, he proved that it would flatten, but not what the shape would be exactly). It wasn't until many years later (1740) that Clairaut and Maclaurin actually proved that an oblate spheroid shape was in fact an exact equilibrium state. They showed that a spheroid with any eccentricity (so-called Maclaurin spheroids) was an equilibrium configuration, given that it had the right amount of angular momentum [23]. Intuitively, most of us would probably not go further than Newton did in assuming that the only equilibrium configurations of such a rotating liquid mass were axisymmetric (symmetric along the axis of rotation) since there would be nothing to break the symmetry. However, it was shown in 1834 by Jacobi that, for certain values of angular momentum greater than some critical value, there exists equilibrium ellipsoids with three unequal axes (so-called Jacobi ellipsoids, see Fig. 1.1).

So far the question being asked was what the possible equilibrium configurations looked like, but not whether they happened in reality or when they happened – after all, there were certain values of angular momentum where both the Maclaurin spheroids and Jacobi ellipsoids were possible equilibrium conditions, so which one would appear in real life? It wasn't until C. O. Meyer and Liouville entered the picture in the mid 1840's that this question was addressed. Their results found that if you go along the Maclaurin sequence of spheroids with increasing angular momentum there is a certain point at which a bifurcation occurs: now any perturbation of the form of a second order harmonic[2] grows (neutrally) and the spheroid becomes a Jacobi ellipsoid. Furthermore, this bifurcation point coincides with the beginning of the Jacobi sequence. Forty or so years later in 1885 Poincaré, with the power of ellipsoidal harmonic analysis,[3] decided

[1] Here "self-gravitating" essentially means that we take into account the gravitational field of the fluid around the point of interest. See Chapter 3 for a precise definition.

[2] Second order ellipsoidal harmonic, to be precise, which basically looks like simultaneous elongation of one axis and shrinking of the the other, both of which are perpendicular to the axis of rotation, which remains fixed.

[3] Poincaré had the good fortune of being in a time when ellipsoidal harmonic analysis had had time to mature, otherwise it's debatable whether what he did would've been possible.



(a) An example of a Maclaurin spheroid, whose axes a_1 and a_2 (corresponding to the x_1 and x_2 directions) are equal and the rotational axis a_3 is the smaller of the three.

(b) An example of a Jacobi ellipsoid, with axes $a_1 = 1, a_2 = 0.66, a_3 = 0.78$ (all normalized by the value of a_1) and a corresponding value of $\Omega = 0.603(Gm/a_1^3)^{1/2}$ rad/s.

FIGURE 1.1: A comparison of a Maclaurin spheroid and a Jacobi ellipsoid.

to tackle the harder question of the stability of the Jacobi ellipsoids. Looking at the the limit where the angular momentum goes to infinity, the Jacobi ellipsoids begin to look more and more like a long baguette - the longest axis grows towards infinity while the smaller two axes (one of which is the axis of rotation) go to zero. Intuitively, this extreme elongation of the Jacobi ellipsoid with increasing angular momentum cannot possibly go on forever, but then when will it become unstable and what happens when it does? Poincaré was able to show that, at a certain value of the angular momentum, even the Jacobi ellipsoids become unstable to ellipsoidal harmonic modes. In this case, the first instability is that of a growing third order harmonic, which gives rise to the famous "pear-shaped" instability shape. The natural question is: what happens next? Is this pear shape stable or unstable? That is, will a new equilibrium sequence branch off from the Jacobi ellipsoids meaning that the pear shape is a legitimate equilibrium family (like what happened with the Jacobi ellipsoids), or is this instability what we call "ordinarily" unstable (i.e. not neutral like the instability causing the bifurcation of the Maclaurin spheroids) and is therefore followed by more complicated dynamics? These questions were tackled by people like Poincaré, Darwin, Liapounoff, Cartan, and Jeans, to name a few. It was Cartan who finally proved that the pear shape was unstable and could not be a branch of equilibrium configurations [4]. Unfortunately, the question of what happens next is then much harder because now one can no longer use ellipsoidal harmonic analysis to investigate stability (since the shape is no longer an ellipsoid). We are constrained to look at the linear motion for small times and velocities or use other, more modern, techniques to study the system. Fortunately, this difficulty did not stop people from being interested in the problem, extending it, and studying its properties.

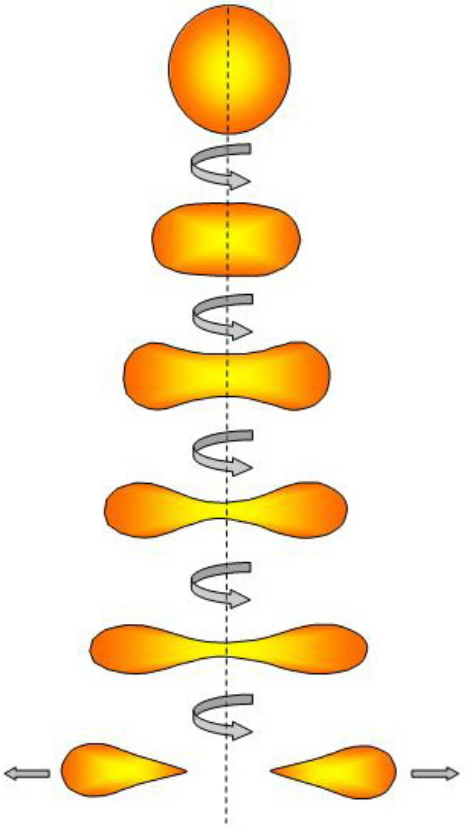


FIGURE 1.2: A sketch of the mechanism of binary star formation given an instability of the Jacobi sequence, according to Darwin's intuition. Source for image: [33]

While for some these problems were simply mathematical curiosities, the series of instabilities and possible equilibrium shapes intrigued Darwin and Poincaré with its possible physical application. Darwin pictured the resulting pear-like instability of the Jacobi ellipsoids as a mechanism for splitting the object in two (see Fig. 1.2), thereby forming two new, star-like objects, and saw this as a possible explanation for the formation of binary stars. Unfortunately, like I mentioned above, this analysis of the resulting dynamics after the pear-like instability was almost unapproachable in a rigorous sense, so it was only speculation from Darwin. However, even if we assume that this instability causes the splitting of the two, it won't give a very satisfactory answer to Darwin's proposition about binary star formation. This is because, if the so-called star (keep in mind this is a pretty crude approximation of a star) splits up, the two pieces will be of different sizes and they will be traveling away from each other towards infinity rather than orbiting one another [23].

Although the promising physical application of this problem to astrophysics does not seem to be very successful, there remain very interesting questions on the subject of studying equilibrium ellipsoids as both a neat mathematical problem and one with other interesting applications. One very fruitful extension to the question of the equilbria and

stability of rotating ellipsoids, with possible applications to modeling nuclei, rotating stars, galaxies, gaseous plasmas, or elastic bodies [5, 8, 10, 18, 22, 29], was first posed by Dirichlet in 1857. He wondered, "under what conditions can one have a configuration which, at every instant, has an ellipsoidal figure and in which the motion, in an inertial frame, is a linear function of the coordinates? [4]" Previously, what had been studied was a self-gravitating liquid mass undergoing solid-body rotation, but now Dirichlet wondered what would happen if there was an inviscid flow whose velocity depended linearly on coordinates in the rotating frame.[4] While Dirichlet formulated the question and began to study it, this seemed to have been simply a side project that he did not spend too much time on, so he was only able to look at spheroids. However, his student, Dedekind, who published Dirichlet's work on this problem posthumously, worked on the full problem including ellipsoids. He is most recognized for proving a very interesting theorem regarding the possible linear dependence of the velocity on the coordinates. I will state his theorem now:

Theorem 1.1 (Dedekind's Theorem)**.** *If a homogeneous ellipsoid with semi-axes a_1, a_2, and a_3 is in gravitational equilibrium, with a prevalent motion $\vec{u}^{(0)}$ (measured in the inertial frame but resolved along the axes rotating along with the ellipsoid) whose components are given by:*

$$\vec{u}^{(0)} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} \begin{pmatrix} x_1/a_1 \\ x_2/a_2 \\ x_3/a_3 \end{pmatrix} = \mathbf{A} \begin{pmatrix} x_1/a_1 \\ x_2/a_2 \\ x_3/a_3 \end{pmatrix}, \tag{1.1}$$

then, the same ellipsoid will also be a figure of equilibrium if the prevalent motion is that derived from the transposed matrix $\mathbf{A^t}$, i. e. for $\vec{u}^{(0)}$ given by

$$\vec{u}^{(0)} = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \alpha_{31} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} \\ \alpha_{13} & \alpha_{23} & \alpha_{33} \end{pmatrix} \begin{pmatrix} x_1/a_1 \\ x_2/a_2 \\ x_3/a_3 \end{pmatrix} = \mathbf{A^t} \begin{pmatrix} x_1/a_1 \\ x_2/a_2 \\ x_3/a_3 \end{pmatrix}. \tag{1.2}$$

[4] Where the configuration with the motion derived from $\mathbf{A^t}$ is called the **adjoint** *of the configuration with the motion derived from $\mathbf{A}$.*

The quest to find the possible equilibrium figures given Dirichlet's assumptions was completed by Riemann in 1861 [28] in a celebrated paper in which he reformulates the problem in an elegant way using a pseudo-potential. In this paper, Riemann was able to determine and classify all the possible relative equilibrium conditions as well as study their stability using an energy criterion including this pseudo-potential. He

[4] This turns out to be equivalent to saying that the flow is of constant vorticity even in the rotating frame.

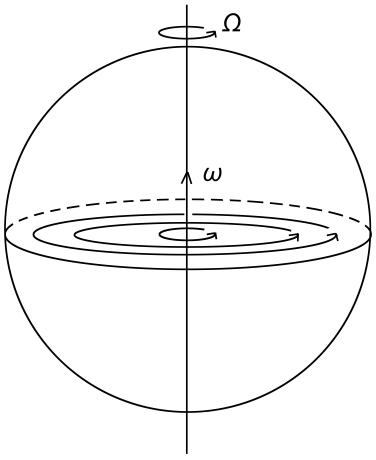


FIGURE 1.3: A schematic drawing of the specific situation we are looking at throughout the rest of this work. We have solid body rotation $\vec{\Omega}$ and internal vorticity $\vec{\omega}$ both aligned with the a_3 axis of the ellipsoid.

found that the most general form of equilibrium is a set of specific superpositions of solid-body rotation $\vec{\Omega}$ and uniform internal vorticity $\vec{\omega}$ in the rotating frame. There are three "types:" no internal motion and just solid-body rotation (which is what was studied in the past by Newton, Maclaurin, Jacobi, and the rest); when $\vec{\Omega}$ and $\vec{\omega}$ coincide with principal axes of the ellipsoid; and when $\vec{\Omega}$ and $\vec{\omega}$ lie on a principal plane of the ellipsoid. *In this work, we will be looking at the incompressible second case and more specifically the case when $\vec{\Omega}$ and $\vec{\omega}$ are aligned with the a_3 principal axis* (see Fig. 1.3), which lies along the x_3 direction, but more on that later (for a complete study of this specific case done by Chandrasekhar, look at [3]). For this case, it turns out that the set of equilibrium ellipsoids is a two-parameter family, with parameters $f := \omega_3/\Omega_3$ and $\tilde{a}_3 := a_3/a_1$. Note that the vorticity and angular velocity vectors are now simply the x_3 components of the vectors because we have assumed them to be aligned with that axis.

Using an energy criterion, Riemann was able to make many claims about the stability of these so-called **"Riemann Ellipsoids"**[5] (although really they should be called "Dirichlet Ellipsoids," but the terminology is not up to me!), all which went undisputed until the 1960's when Chandrasekhar and Norm Lebovitz published a series of papers where they did a very thorough study of the problem (these works were combined and released as a book by Chandrasekhar in 1969 [4]). Lebovitz and Chandrasekhar approached the study of the stability in a different way than had been previously done, thanks to the novel application of the tensor form of the virial equation used to reduce the degrees of freedom of the system in question, taking advantage of symmetries and relatively simple structure of the flow and shape of the object. This so-called "virial method"

[5] To avoid unnecessary complications, I will not go into the details of these claims and will only focus on our specific situation which we are studying.

(developed for this problem by Lebovitz [16]) essentially comprises of multiplying the jth component of the momentum equation by x_i then integrating, resulting in a system of ordinary differential equations, which significantly reduces the complexity of the problem (this is the second-order virial method, more specifically). This reduction of degrees of freedom is in fact *exact* (that is, it's not an approximation of the dynamics) due to the fact that the ellipsoid remains an ellipsoid – consequently, this means that it is only good for studying perturbations which maintain the ellipsoidal shape, i.e. second-order harmonics. Note that we can look at the *third* order virial by multiplying the kth component of the momentum equation by $x_i x_j$ and integrating. One can use this to study perturbations of the form of third order harmonics (which is in fact what is done in [3], for example). Using this technique, Chandrasekhar and Lebovitz looked at the spectral stability of Riemann Ellipsoids[6]. Their results disagree with that of Riemann's for certain classes of ellipsoids[7], including the case which we are studying. Riemann's criterion predicts that all ellipsoids lying between $f = -\frac{a_1^2+a_2^2}{a_1 a_2}$ and $f = -2$, and for all values of $\tilde{a}_3$, are *unstable*, whereas Chandrasekhar predicts regions within that set of parameters which contain spectrally stable configurations. Now, spectral instability implies nonlinear (or Lyapunov) instability, but not necessarily the other way around – so it is possible that the regions of spectral stability are still non-linearly unstable, which would not necessarily be a contradiction to Riemann's results. However recent computational results show that this spectrally stable region predicted by Chandrasekhar is what is called "Nekhoroshev stable," which is essentially a weaker version of Lyapunov stable (it computationally looks at the distance of a point after long time, rather than infinite time) [11]. So where did Riemann go wrong? Typically, the statement of the energy criterion for stability of a system consists of two claims: 1) If the potential energy is a minimum at equilibrium, then the equilibrium configuration is stable (what is called "Lyapunov stable"). 2) If it is *not* a minimum, then it is *unstable*. So normally this is an "if and only if" condition for the stability of a system. Notice I said normally – it turns out that if the system is incapable of being put into Lagrangian form *one can only conclude statement 1) but can't say anything about the system being unstable if the potential energy is not a minimum.* Lebovitz proved that Riemann's system cannot be put into a Lagrangian form, thereby putting into question Riemann's conclusions of instability (which is where the results disagree) [17].

Although many studies of Riemann Ellipsoids were done in the 1960's through the 1980's, including possible applications, generalizations to compressible fluids and plasmas, and so on, we will close out this section by mentioning the Hamiltonian formulation of

[6] An important thing to note is that, clearly instability in the reduced system implies instability in the full (non-reduced) system, but the same cannot be said about the *stable* case.

[7] Please refer to [17] for an extensive explanation of the discrepancies between Riemann and Chandrasekhar's results, including an explanation for Riemann's process and possible error.

the problem, which was originally attempted by Rosensteel in 1988 [29] and completed by P. J. Morrison, Lebovitz, and J. Biello in 2009 [27] (see Chapter 2 for a complete description). While Rosensteel was able to successfully put the problem in terms of a Hamiltonian and a Poisson bracket, he did not include the incompressibility constraint in his formulation, leading to a non-physical Hamiltonian. However, Morrison et al. correctly formulated the problem by applying Dirac constraints to the full fluid dynamics Poisson bracket [25], then applying the virial method to the Poisson bracket to get a modified (constrained) "Rosensteel bracket" which, when fed the *physical* Hamiltonian, gave the correct equations of motion. This Hamiltonian formulation has the advantage of giving us better tools to analyze the stability of the Riemann Ellipsoids – it makes the spectral stability problem far less messy and it also provides more reliable and stronger stability results, such as Dirichlet's Theorem [24]:

Theorem 1.2 (Dirichlet)**.** *Let the second variation of the Hamiltonian $\delta^2 H$ be definite at an equilibrium z^e. Then z^e is stable.*

However, for a non-canonical Hamiltonian system, such as ours, this theorem does not hold and we are forced to use a similar theorem which applies to such systems using what is called the "Energy-Casimir." The theorem states [24]:

Theorem 1.3. *Let the Hessian matrix $\frac{\partial^2 F}{\partial z^\alpha \partial z^\beta}$ of the Energy-Casimir F be definite at an equilibrium z^e. Then z^e is stable.*

If this holds for our family of ellipsoids, then we can prove that they ellipsoids are in fact stable in the nonlinear sense. Although the Hamiltonian formulation has been done, a stability analysis of this problem using the Hamiltonian formulation has yet to be completed, and this is in fact one of the main goals of the thesis.

1.2 Summary and Outlook

Quoting Chandrasekhar from [3], "The problem that is to be considered in this paper [and in this thesis] is that of a homogeneous mass, rotating uniformly with an angular velocity $\vec{\Omega}$, with internal motions having a uniform vorticity $\vec{\omega}$ in the direction of $\vec{\Omega}$," and this direction being the x_3 direction, along the a_3 principal axis of the ellipsoid. Furthermore, we will be looking at a case where the fluid is incompressible ($\nabla \cdot \vec{U} = 0$) and where viscosity is zero. Like I mentioned above, in this situation it turns out that the set of equilibrium ellipsoids is a two-parameter family, with parameters $f := \omega_3/\Omega_3$ and $\tilde{a}_3 := a_3/a_1$. We will focus solely on studying the stability of the ellipsoids lying

on the so-called "self-adjoint sequence[8]," meaning that $f = -\frac{a_1^2+a_2^2}{a_1 a_2}$, by looking at the stability as we change $\tilde{a}_3$.

In Chapter 2 we will go over the necessary mathematical, Hamiltonian, and fluid dynamics background to do the Hamiltonian formulation given by Morrison et al., including Dirac constraints, where I will also present the novel approach to this formulation which is part of my work. This formulation is the one that we will be utilizing to do stability analysis throughout this work. Then in Chapter 3 we will look at the classical formulation of the problem given by Dirichlet and the virial form given by Chandrasekhar (it will be a somewhat hybrid formulation, making it a bit neater and normalizing notation), where we look at the equations of motion for the virial tensors. In Chapter 4 I will go over my derivation of the equilibrium conditions given our specific value of f, then go over my own stability analysis using the Hamiltonian formulation followed by Chandrasekhar's spectral stability analysis and a comparison. Finally, in the last section, I will go over the results and discuss their implications, ending with a conclusion of my work in Chapter 5.

[8]Self-adjoint here meaning that its adjoint flow (in the context of Dedekin's theorem, Theorem 1.1) is equal to itself.

Chapter 2

Hamiltonian Formulation

Studying the Hamiltonian formulation of our problem requires a precise mathematical toolset that allows us to describe Hamiltonian systems and look at their properties. The advantage of Hamiltonian systems is that they are described by the very elegant mathematics of symplectic topology, which not only provides a nice geometric (and therefore more intuitive) picture of many problems, but also allows us to study very general properties that apply to many systems of varying dimensions and characteristics.

This chapter is devoted to generalizing much of what we learn about Hamiltonian dynamics in classes via a very mathematically precise approach, and then applying this theory to our specific problem. For most of the first section I will be giving the necessary mathematical background for understanding Hamiltonian systems and their properties. I will be using various sources throughout this chapter, including *Guillemin & Pollack* [12], *Lee* [21], Dr. Andrew Neitzke's notes from his differential topology class, *da Silva* [9], *Arnold* [1], *Morrison* [24], and *Audin* [2]. In a situation in which I feel it is unnecessary to divulge all of the details of, say, a proof, I will give further information on where to find a complete version. Finally, this section assumes basic knowledge of linear algebra, calculus in Euclidean space, real analysis, and point-set topology[1].

2.1 Mathematics Background

2.1.1 Differential Topology

I will begin by developing some basic tools of differential topology – the study of smooth manifolds. J. M. Lee's book on smooth manifolds opens with the following quote [21]:

[1] Munkre's *Topology* is a good source for those who need to build a basis in point-set topology before moving on.

> Manifolds crop up everywhere in mathematics. These generalizations of curves and surfaces to arbitrarily many dimensions provide the mathematical context for understanding "space" in all of its manifestations.

It is no surprise, then, that manifolds crop up everywhere in *physics* as well, considering that the point of physics is to explain the properties of matter and energy living in *space*. Furthermore, this space is not always the same and it is not always easy to work with; for example, we live on a sphere (the Earth), space-time is curved, and the state-space of fluid dynamics is an infinite-dimensional function space. Having the tools to describe these spaces, and therefore the dynamics within them, is an essential part of physics and why I am devoting a whole section of this chapter to developing the mathematical tools to do so.

We begin by defining a *topological manifold*, which, intuitively, is some topological space which locally looks like $\mathbb{R}^n$[2].

Definition 2.1 (Topological Manifold)**.** Let M be a topological space. Then M is a **topological n-manifold** (of dimension n) if:

1. M is a *Hausdorff space*: for every pair of distinct points $p, q \in M$, there are disjoint open subsets $U, V \subseteq M$ such that $p \in U$ and $q \in V$.
2. M is *second countable*: there exists a countable basis for the topology of M.
3. M is *locally Euclidean of dimension n*: for all $p \in M$, there exists an open neighborhood $U \subset M$ (containing p) and a homeomorphism $\phi : U \to V$, with $V \subset \mathbb{R}^n$ open.

At this point, it is still not very clear why this definition helps us do anything on this space. The idea behind studying manifolds is that we know how to do things in $\mathbb{R}^n$, so we can use the fact that it looks like $\mathbb{R}^n$ to make our lives easier. We begin by defining a *coordinate chart* on M, which will be essentially the tools to navigate our way around the topological manifold M using our familiar space of $\mathbb{R}^n$.

Definition 2.2 (Coordinate Chart)**.** Let M be a topological n-manifold, U an open subset of M, and $\phi : U \to V \subset \mathbb{R}^n$ a homeomorphism. A **coordinate chart** on M is the pair (U, ϕ).

Since M is a topological manifold, such a chart (U, ϕ) is guaranteed to exist around every point $p \in M$. Therefore, we can cover all of M by such charts, and this collection

[2]For the sake of brevity I will not use generic affine spaces $\mathbb{A}^n$, since everybody is familiar with all the vector space properties of $\mathbb{R}^n$.

is called an **atlas for** M. In fact, because every open cover of a second-countable space has a countable sub cover, M is covered by *countably* many charts $\{(U_i, \phi_i)\}$.

We can view this coordinate chart as being what physicists normally consider to be "local coordinates:" a grid which is spread out over some patch of the manifold M which gives us an n-tuple of numerical values (coordinates) for each point in $U \subset M$. See Figure 2.1 for a picture of what I am describing. In a lot of situations, the domain of

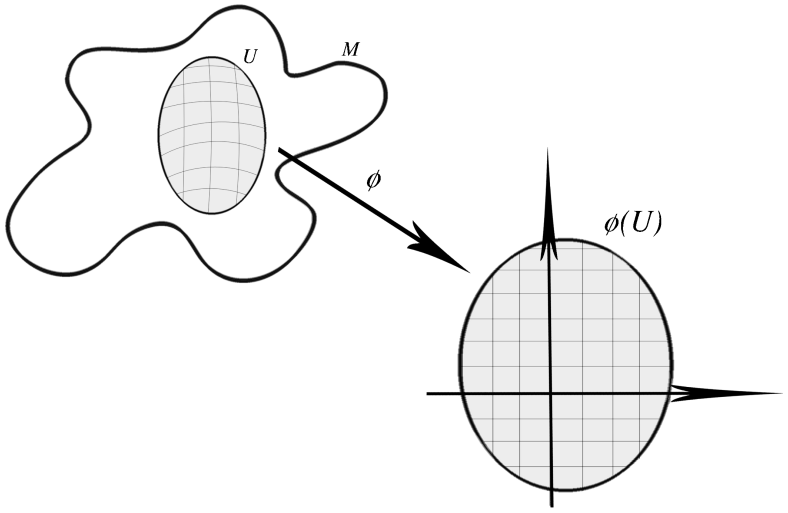


FIGURE 2.1: The intuitive picture of how a coordinate chart (U, ϕ) can be seen as a "local coordinate grid" on M, which gives a local representation for a point $p \in U$, say.

these charts may overlap on M. When two chart domains $U_\alpha, U_\beta \subset M$ overlap, we can define a function $\tau_{\alpha\beta} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$, called a **transition function** (or transition map), which tells us how to go from one coordinate system to another on M. These functions are maps from $\mathbb{R}^n$ to itself, so we can use all of the tools from calculus and real analysis to study these functions. But first a little bit of intuition and general understanding.

Now, usually I tend to intuitively picture manifolds as begin embedded in some higher dimensional space, so that I can picture them as two-dimensional surfaces in a three-dimensional space. Although it is possible to embed an arbitrary n-manifold into $\mathbb{R}^m$ for some m, this picture is not always very helpful and it may be detrimental, even. Another way to view an n-manifold is to see it as collection of subsets of $\mathbb{R}^n$ (the collection $\{\phi_\alpha(U_\alpha)\}_\alpha$) and functions ($\{\tau_{\alpha\beta}\}_{\alpha\beta}$) "gluing" them together! Here is a helpful analogy: if one wants to travel around the world, you can look at a 3D globe, but this may not be a very convenient thing to have if you were on a road-trip, say. However, what one could do is use a road map/atlas. Now instead of having a 3D globe, you have a collection of pages (a collection of 2D subsets) which tell you where to go. But this is not all the information they carry – what if you are traveling so far that you go beyond the scope of

one page? Then what you must do is look at the overlap region of your page to the next page (which you will be traveling into) and use the atlas to find yourself in the other page (i.e. another coordinate chart). Changing page is the equivalent to doing a coordinate change with $\tau_{\alpha\beta}$. According to these gluings, you can have a variety of many different topological manifolds of dimension 2, as in the analogy. This gluing is essential! One simple example to understand this importance is that of a cylinder and a Möbius band. These two 2-manifolds can be covered with two charts, which represent two copies of some open subset of $\mathbb{R}^2$. Without the gluing information, one can't distinguish between the two. However, going from one chart to the next, one finds in the cylinder that the chart is just a rotated version of the other one, whereas the Möbius band requires a twist in the gluing when going from one overlap to the other. As we are about to find out, there is another important property that we can look for in these transition (gluing) maps and it is going to be essential for our next step.

Definition 2.3. Let M be a topological manifold, and $(U_\alpha, \phi_\alpha), (U_\beta, \phi_\beta)$ two charts on M. Then (U_α, ϕ_α) and (U_β, ϕ_β) are **C^∞-related** (or smoothly compatible) if $\tau_{\alpha\beta}$ is a diffeomorphism (in the usual $\mathbb{R}^n$ sense).

If two chart domains don't overlap, then they're trivially C^∞-related. This relation allows us to define another important idea:

Definition 2.4 (Smooth Atlas)**.** Let M be a topological n-manifold and $\mathcal{A}$ the index for the charts covering M. Then a **smooth atlas** on M is a collection $\mathcal{C} = \{(U_\alpha, \phi_\alpha)\}_{\alpha\in\mathcal{A}}$ of charts such that

1. $\bigcup_{\alpha\in\mathcal{A}} U_\alpha = M$.
2. If $\alpha_1, \alpha_2 \in \mathcal{A}$ then $(U_{\alpha_1}, \phi_{\alpha_1})$ and $(U_{\alpha_2}, \phi_{\alpha_2})$ are C^∞-related.

Given any smooth atlas $\mathcal{C}$, we can uniquely extend $\mathcal{C}$ to a smooth *maximal* (with respect to 1 and 2) atlas of M, called a **smooth structure on M**. In practice, one only needs to come up with a smooth atlas of M and a smooth structure given by its extension is well-defined and unique on M. For a proof of this claim, see Proposition 1.17 in Lee's book [21]. We are finally ready to define a smooth manifold.

Definition 2.5 (Smooth Manifold)**.** Let M be a topological n-manifold and $\mathcal{C}$ be a smooth structure of M. Then the pair $(M, \mathcal{C})$ is a smooth manifold.

Topological manifolds are nice because they locally look like some Euclidean space, that is, they're homeomorphic to some subset of Euclidean space. However a homeomorphism doesn't guarantee that the space doesn't have "creases" or "wrinkles" (the intuition

being in the context of a 2D surface embedded in 3D space). We want just a little bit more niceness –to solve equations of motion requires calculus, which, as we will see, cannot be done on surfaces which are creased or wrinkled (i.e. *not* smooth). Smooth manifolds allow us to expand the idea of calculus, which we are used to doing in Euclidean space, to more general topological spaces – namely, smooth manifolds.

Now that we have a definition of a smooth manifold, we can talk about smooth *functions* between smooth manifolds.

Definition 2.6. Let M, N be smooth manifolds, and $f : M \to N$ a function between the two. Then we say that f **is smooth at** $p \in M$ if, for all charts (U_α, ϕ_α) on M containing p and all (V_β, ψ_β) on N containing $f(p)$, $\psi_\beta \circ f \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap f^{-1}(V_\beta)) \to \psi_\beta(V_\beta)$ is smooth at $\phi_\alpha(p)$, in the usual $\mathbb{R}^n$ sense.

Once again, it is not necessary to check smoothness for *every* chart domain containing p. In fact, f is smooth at $p \in M$ if and only if there exist single charts (U_α, ϕ_α) of M and (V_β, ψ_β) of N such that $\psi_\beta \circ f \circ \phi_\alpha^{-1}$ is smooth at $\phi_\alpha(p)$. This is because any two charts in the smooth atlases of M and N are C^∞-related, so changing coordinates is a smooth thing, and the composition of smooth functions is smooth. A function f is said to be **smooth** if it is smooth at p for every $p \in M$, and a **diffeomorphism** if it is smooth and bijective, with a smooth inverse. Note that this definition depends on the smooth structure of M, so if two smooth atlases give two different smooth structures on M, then a function f may be smooth in one and not smooth in the other. One classic example of having two different smooth structures is that of the smooth structure pertaining to the topological space $\mathbb{R}$. The atlas $(\mathbb{R}, id_\mathbb{R})$ gives a smooth structure on $\mathbb{R}$, but $(\mathbb{R}, \psi)$, with $\psi(x) = x^3$, gives another. This is because the two charts are not C^∞-related (and hence not in each others smooth maximal atlas). At the origin, ψ is not smooth in the usual sense that we know for $\mathbb{R}$, and so it is not smoothly related to the chart $(\mathbb{R}, id_\mathbb{R})$.

Just above, we were looking at a map $f : M \to N$, which maps points in M to points in N, and discussed a definition of smoothness for f in terms of derivatives of its coordinate representation $\psi_\beta \circ f \circ \phi_\alpha^{-1}$. How do we translate the derivatives of $\psi_\beta \circ f \circ \phi_\alpha^{-1}$ to some general notion of the derivative of f, which is coordinate-free? Well, in Euclidean space one can see the derivative matrix of a function $h : \mathbb{R}^n \to \mathbb{R}^m$, $\frac{\partial h^i}{\partial x^j}$ as a linear function acting on *vectors* (not points! Using affine space would have been useful to resolve this ambiguity) in $\mathbb{R}^n$ and sending them to vectors in $\mathbb{R}^m$. So from h we have defined a map from the corresponding vector space of the domain to the vector space associated with the range. We can do something similar for a function of manifolds $f : M \to N$, except now there is no clear vector space attached at each point like there was in $\mathbb{R}^n$, so we have to formalize this using charts, which gives us a connection from our manifold to

some affine space with an overlying vector space. Let M be an m-dimensional smooth manifold and $p \in M$. Then we are guaranteed a chart (U_α, ϕ_α) around p which sends U_α to some open subset of $\mathbb{R}^m$. Now, like I mentioned above, the point $\phi_\alpha(p) \in \mathbb{R}^m$ has an obvious vector space "attached" to it at that point, namely another copy of $\mathbb{R}^m$, so we can choose a vector $\xi \in \mathbb{R}^m$ anchored at $\phi_\alpha(p)$ and associate it with the point $p \in M$. The only problem is that this vector depends on the chart that we have chosen, but we want to define a vector at $p \in M$, which should not depend on a chart choice. This brings us to the definition of a *tangent space*:

Definition 2.7 (Tangent Space)**.** Let M be a smooth m-manifold with atlas $\mathcal{C} = \{(U_\alpha, \phi_\alpha)\}_{\alpha \in \mathcal{A}}$, $p \in M$. Then the **tangent space of** M **at** p is defined to be

$$T_pM = \{(\alpha, \xi) \mid \alpha \in \mathcal{A}, p \in U_\alpha, \xi \in \mathbb{R}^m\}/\sim, \tag{2.1}$$

where $(\beta, \xi) \sim (\alpha, d(\phi_\alpha \circ \phi_\beta^{-1})_{\phi_\beta(p)}(\xi))$.

Given a chart U_α around $p \in M$, we have a canonical isomorphism: $j_\alpha : T_pM \to \mathbb{R}^m$, given by $j_\alpha((\alpha, \xi)) = \xi$. If we have another representation, then it's $j_\alpha((\beta, \eta)) = j_\alpha((\alpha, d(\phi_\alpha \circ \phi_\beta^{-1})_{\phi_\beta(p)}(\eta))) = d(\phi_\alpha \circ \phi_\beta^{-1})_{\phi_\beta(p)}(\eta)$. It is not hard to show that this definition of tangent space produces a vector space, but for the sake of brevity, I will not do so here. In other words, a tangent vector of M at the point p can be *represented* by a vector $\xi \in \mathbb{R}^m$ of some chart (U_α, ϕ_α), and called (α, ξ), but also by any vector $\eta \in \mathbb{R}^m$, for (U_β, ϕ_β) another chart, as long as $d(\phi_\beta \circ \phi_\alpha^{-1})(\xi) = \eta$. See Figure 2.2 for a bit of elaboration with pretty pictures (that I worked so hard to make!). Of course, in certain situations we have what we always picture as tangent spaces – a plane grazing a 2-manifold in $\mathbb{R}^3$ with vectors tangent to the surface living in $\mathbb{R}^3$. This is the most intuitive picture, although somewhat misleading in most cases when there's no obvious embedding into a higher-dimensional $\mathbb{R}^k$.

Okay, now that we have generalized what it means to have a vector attached to each point of the manifold we can go back and discuss the idea of a differential of f at $p \in M$, which we will call df_p, mapping vectors in T_pM to vectors on $T_{f(p)}N$. In fact, let's define:

Definition 2.8. Let M, N, be smooth m- and n-manifolds, respectively, $f : M \to N$ a smooth function. Then define $df_p : T_pM \to T_{f(p)}N$ by

$$df_p = j_\beta^{-1} \circ d(\phi_\beta \circ f \circ \phi_\alpha^{-1}) \circ j_\alpha, \tag{2.2}$$

for any chart (U_α, ϕ_α) of M, such that $p \in U_\alpha$, and (V_β, ϕ_β) of N, such that $f(p) \in V_\beta$.

We must check that this definition is well-defined and doesn't depend on charts chosen. Let's change charts, denoting the new chart on M by index α' and the one on N by β'.

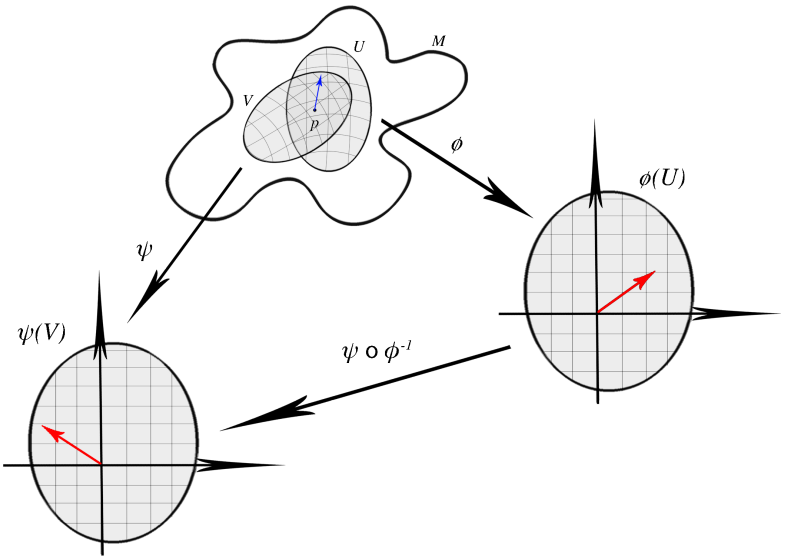


FIGURE 2.2: Here is an image depicting two representations (in red) of a tangent vector (in blue). One could picture the tangent space as being a collect of $\mathbb{R}^n$ with "gluings" defined by $d\tau_{\alpha\beta}$, for each pair of charts (U_α, ϕ_α) and (U_β, ϕ_β) in $\mathcal{C}$, which is very similar to the picture we had for manifolds.

We have:

$$\begin{aligned} df_p &= j_{\beta'}^{-1} \circ d(\phi_{\beta'} \circ f \circ \phi_{\alpha'}^{-1}) \circ j_{\alpha'}, \\ &= j_\beta^{-1} \circ j_\beta \circ j_{\beta'}^{-1} \circ d(\phi_{\beta'} \circ f \circ \phi_{\alpha'}^{-1}) \circ j_{\alpha'} \circ j_\alpha^{-1} \circ j_\alpha, \\ &= j_\beta^{-1} \circ d(\phi_\beta \circ \phi_{\beta'}^{-1}) \circ d(\phi_{\beta'} \circ f \circ \phi_{\alpha'}^{-1}) \circ d(\phi_{\alpha'} \circ \phi_\alpha^{-1}) \circ j_\alpha, \\ &= j_\beta^{-1} \circ d(\phi_\beta \circ \phi_{\beta'}^{-1} \circ \phi_{\beta'} \circ f \circ \phi_{\alpha'}^{-1} \circ \phi_{\alpha'} \circ \phi_\alpha^{-1}) \circ j_\alpha, \\ &= j_\beta^{-1} \circ d(\phi_\beta \circ f \circ \phi_\alpha^{-1}) \circ j_\alpha. \end{aligned}$$

And so, we see that this is indeed a well-defined, coordinate-independent map.

As you saw in Definition 2.7, every point $p \in M$ has a tangent space, and it turns out to be useful to consider all of these tangent spaces as one object, called the *tangent bundle of* M, TM. We define the **tangent bundle of** M to be the disjoint union of all the tangent spaces of the manifold M, i.e.,

$$TM = \bigsqcup_{p \in M} T_pM. \tag{2.3}$$

An element of this tangent bundle TM is written as a pair (p, v) by identifying a point $p \in M$ and a vector $v \in T_pM$ which is tangent to the point. Furthermore, for any element of TM, there exists a projection map $\pi : TM \to M$ which gives us the base point on M, so $\pi((p, v)) = p$. It may seem, by the way it's been presented, that the tangent bundle is just the product $M \times \mathbb{R}^n$, but this is not necessarily true. In fact, it

is only *locally* true for every tangent bundle (and true globally for some specific tangent bundles). For a chart (U_α, ϕ_α) of M, we can define what is called a *local trivialization* by defining: $(\phi_\alpha \times j_\alpha^{-1})(p, v) = (\phi_\alpha(p), j_\alpha^{-1}(v))$. This map, which is only guaranteed to exist *locally* gives us a homeomorphism from $TM|_{U_\alpha}$ to $\phi_\alpha(U_\alpha) \times \mathbb{R}^m$. The global topology of TM as a smooth manifold[3], similar to before with manifolds, is determined by the transition functions $\tau_{\alpha\beta} \times d(\tau_{\alpha\beta})$. Finally, we define:

Definition 2.9 (Smooth Section)**.** Let M be a smooth m-manifold with tangent bundle TM and projection $\pi : TM \to M$. Then a **smooth section** of TM is the map $s : M \to TM$ such that $\pi \circ s = id_M$. Furthermore, define $\Gamma(TM)$ to be the space of possible smooth sections of TM.

Every tangent space T_pM is a vector space, as we have seen, and so naturally there exists a dual to this tangent space, called the **cotangent space** $(T_pM)^* := T_p^*M$. Similarly, one can define the **cotangent bundle** T^*M by

$$T^*M = \bigsqcup_{p \in M} T_p^*M, \tag{2.4}$$

where now the local trivialization $\phi_\alpha \times j_\alpha^*$ is defined such that $j_\alpha^*(\eta) \cdot j_\alpha(v) = \eta(v)$, where $(\ \cdot\ )$ is the dot product in $\mathbb{R}^m$, $\eta \in T_p^*M$, and $v \in T_pM$. In other words, there is a natural pairing of η with v and the trivialization on T^*M respects this pairing at each point. One last thing to point out about the cotangent bundle is the basis induced by the local trivialization at the cotangent space at a point p. Let (U_α, ϕ_α) be a chart on M, then we saw earlier that $j_\alpha(p) : T_pM \to \mathbb{R}^m$ is an isomorphism which we can use to define $\partial/\partial\phi_\alpha^i$ as $[j_\alpha^{-1}(\phi_\alpha(p))](e_i)$, where e_i is the ith basis vector in $\mathbb{R}^m$. Then $\{\frac{\partial}{\partial\phi_\alpha^i}(p)\}_i$ forms a basis for T_pM at each $p \in U_\alpha$. Now, we can do something similar for the cotangent bundle by defining $d\phi_\alpha^i$ by $[(j_\alpha^*)^{-1}(\phi_\alpha(p))](e_i)$. A quick check shows us that these are in fact dual basis elements to $\partial/\partial\phi_\alpha^j$: by definition, $d\phi_\alpha^i(\partial/\partial\phi_\alpha^j) = j_\alpha^*(d\phi_\alpha^i) \cdot j_\alpha(\partial/\partial\phi_\alpha^j)$, but with our definitions that we gave last sentence, we have: $d\phi_\alpha^i(\partial/\partial\phi_\alpha^j) = [(j_\alpha^*)^{-1}(\phi_\alpha(p))](e_i)\left([j_\alpha^{-1}(\phi_\alpha(p))](e_j)\right) = j_\alpha^*([(j_\alpha^*)^{-1}(\phi_\alpha(p))](e_i)) \cdot j_\alpha([j_\alpha^{-1}(\phi_\alpha(p))](e_j)) = j_\alpha^*((j_\alpha^*)^{-1}(e_i)) \cdot j_\alpha(j_\alpha^{-1}(e_j)) = e_i \cdot e_j = \delta_{ij}$, which is what we wanted to show.

Now we have built the underlying mathematical framework on which our dynamics will happen. The idea of dynamics implies some sort of "flow" of particles or fluid, say. Therefore, it will be convenient to formalize this idea in precise mathematical terms, which we will then relate to the dynamics I am speaking of (see Subsection 2.2.1 for full details). We begin with the following definition:

[3] I have not proven this, but it is in fact true. See Lee [21] for a proof.

Definition 2.10 (Flow)**.** Let M be a smooth m-manifold. A **flow on** M is a one-parameter family of diffeomorphisms $g_t : M \to M$, $t \in \mathbb{R}$, with the property that $g_t \circ g_s = g_{t+s}$, such that

1. $g : M \times \mathbb{R} \to M$, given by $(p, t) \mapsto g_t(p)$ is smooth.
2. $g_0 = id_M$.

The intuitive idea of velocity of a flow can also be formalized: given a flow g, the **velocity field of** g is the vector field $X \in \Gamma(TM)$ given by

$$X(p) = \left.\frac{d}{dt}\right|_{t=0} g_t(p) = dg_{(p,0)}\left(\frac{\partial}{\partial t}\right), \tag{2.5}$$

with $\partial/\partial t \in T_0\mathbb{R} \subset T_{(p,0)}(M \times \mathbb{R})$. Now, unfortunately, in many physical theories, as is the case for Hamiltonian systems, we are attempting to solve for these flows *given* some velocity field. This is a bit tricky, because there is nothing telling us that this is even possible in any general situation. Thankfully, there's an app – I mean theorem – for that:

Theorem 2.11. *Let M be a compact smooth manifold. Then, given any smooth vector field $X \in \Gamma(TM)$, there exists a flow $g : M \times \mathbb{R} \to M$ with velocity X.*

Proof. M is compact, so a covering by charts (U_α, ϕ_α) can be reduced to a covering by a finite number of charts (U_i, ϕ_i), for $i = 1, \ldots, k$. Looking at an individual chart, we can map the vector field on U_i to a vector field on $\phi_i(U_i) \subset \mathbb{R}^m$. Then we can use the local existence of flows (a result of ordinary differential equations) on open subsets of $\mathbb{R}^m$ for short times, say ϵ_i. This gives us a flow $h_i : \phi_i(U_i) \times (0, \epsilon_i) \to \phi_i(U_i)$, which we can then map to U_i using ϕ_i to get a flow on an open subset of M. Since there are a finite number of charts, we can find an ϵ_{min} such that these local flows all exists, and so a flow exists *globally*, for $t < \epsilon_{min}$. This gives us the existence of $g_{\epsilon_{min}}$, but, using the group structure, we can get a flow for arbitrary time by multiplying: $g_{2\epsilon_{min}} = g_{\epsilon_{min}} \circ g_{\epsilon_{min}}$. This finally gives us $g : M \times \mathbb{R} \to M$. □

It seems as if this theorem is used (or at least assumed to apply) in most situations for Hamiltonian dynamics, since Hamilton's equations really give us a vector field in phase space.

I will now move on to another important aspect of differential topology which we will use later in this chapter: differential forms on M. Defining what a differential form on M is requires us to back up a bit and define some more general structures, namely an exterior algebra. I will do some rapid-fire definitions to get through the basics. We

are looking to define and formalize the idea of differential forms on our manifold M, which, roughly speaking, are objects which, at each $p \in M$, take in k vectors of T_pM and spit out a real number. Said another way, globally, they are objects which take in k vector fields and spit out a smooth function on M, $f \in C^\infty(M)$. These forms will be essential to our theory for various reasons, but they are absolutely essential for defining the equations of motion, and so I think it is worth going into the details.

These differential forms will be constructed, essentially from elements of the cotangent bundle, which take in a vector field and spit out a function. We want the possibility to generalize this for more than one vector field, so we have to understand how to combine these objects in a rigorous way.

Definition 2.12 (Algebra and more.)**.** I will give three definitions.

1. An **algebra** is a vector space A with an associative, linear product $\cdot : A \times A \to A$ and unit $1 \in A$.
2. An algebra A is **$\mathbb{Z}$-graded** if $A = \bigoplus_{n=-\infty}^{\infty} A_n$, where each A_n is a *homogeneous elements of degree* n, and $A_n \cdot A_m \subset A_{n+m}$.
3. A $\mathbb{Z}$-graded algebra A is **graded commutative** if $x \cdot y = (-1)^{|x||y|} y \cdot x$, for x, y homogeneous of degree $|x|, |y|$, respectively.

Given a vector space V, we can construct an algebra by using the tensor product $\otimes$, called a **tensor algebra** $T(V)$, defined to be

$$T(V) = \bigoplus_{n=0}^{\infty} V^{\otimes n} = \mathbb{R} \oplus V \oplus V \otimes V \oplus \cdots \tag{2.6}$$

where the product of the algebra is simply the tensor product, taking a homogeneous element of degree n to one of degree $m + n$ if multiplied by a homogeneous element of degree m. Take for example $v \otimes w \in V^{\otimes 2}$ and $x \otimes y \otimes z \in V^{\otimes 3}$, then $(v \otimes w) \otimes (x \otimes y \otimes z) = v \otimes w \otimes x \otimes y \otimes z \in V^{\otimes 5}$. Note, also, that not every element in $V^{\otimes n}$ is of the form $v_1 \otimes \cdots \otimes v_n$, but is a sum of elements of this form. In fact, if $\{e_1, \ldots, e_k\}$ is a basis for V, the basis elements of $V^{\otimes n}$ are of the form $e_{i_1} \otimes \cdots \otimes e_{i_n}$. One can think of $T(V)$ as analogous to a free algebra, where you haven't said much about the properties, but can map it to an algebra A with structure. But what if we add a little bit more structure to $T(V)$?

Definition 2.13 (Exterior Algebra)**.** Let V be a vector space. Then an **exterior algebra on** V is

$$\wedge(V) = T(V)/I, \tag{2.7}$$

where I is the two-sided ideal generated by $x \otimes x$, with $x \in V$. We write the product of $\wedge(V)$ as $\wedge : \wedge(V) \to \wedge(V)$.

This is not the typical way that one sees exterior algebras defined (at least for physicists), but hopefully the following properties reveal some more practical structure to these exterior algebras. Let $x, y \in V$, then $(x+y) \wedge (x+y) = 0$, since we are modding out by elements of this kind. This implies that $x \wedge y + y \wedge x = 0$ (again, making repeated terms equal 0), which tell us that $x \wedge y = -y \wedge x$. This means that $\wedge(V)$ is a *graded* commutative algebra (enough to check on decomposable elements since the product is linear), where now each homogeneous component $\wedge(V)^n$ has basis $\{e_{i_1} \wedge \cdots \wedge e_{i_n} \mid i_1 < i_2 < \cdots < i_n\}$, and the direct sum of homogeneous components goes up to the dimension of V, i.e.

$$\wedge(V) = \bigoplus_{n=0}^{\dim V} \wedge(V)^n, \tag{2.8}$$

because any basis element of $\wedge(V)^{(\dim V)+1}$ is going to have a repeated basis vector of V and therefore its wedge with itself will be zero. One important thing to note is that just because we are modding out by elements generated by $x \wedge x$, for $x \in V$, does *not* mean that $w \wedge w = 0$ for some $w \in \wedge(V)^n$, with $n > 1$. It is only true for elements which are decomposable, so of the form $w' = e_{i_1} \wedge \cdots \wedge e_{i_n}$, but *not* necessarily true for elements w which are sums of decomposable elements.

So far we have been rather general, but we are looking to work with differential forms acting on vector fields on a manifold M. Therefore, it is natural to consider the exterior algebra of the cotangent bundle, which will be our final goal. First, however, we must generalize this concept of exterior algebras to work on manifolds. This is rather simple: the idea is that you work pointwise, and define and exterior algebra of T_p^*M for each $p \in M$. So, define

$$\wedge(T^*M) = \wedge(T^*M)^0 \oplus \wedge(T^*M)^1 \oplus \cdots \oplus \wedge(T^*M)^m, \tag{2.9}$$

where $\wedge(T^*M)^0 \simeq M \times \mathbb{R}$, $\wedge(T^*M)^1 \simeq T^*M$, and $(\wedge(T^*M))_p = \wedge(T_p^*M)$. Using a very similar definition for $\wedge(TM)$, we define a pairing $\wedge(T^*M)^k \times \wedge(TM)^k \to \mathbb{R}$ given by the pairing of dual elements at each point (again, it is sufficient to define for homogeneous elements)

$$\langle \alpha_1(p) \wedge \cdots \wedge \alpha_k(p), v_1(p) \wedge \cdots \wedge v_k(p) \rangle = \det(\{\alpha_i(v_j)\}_{i,j=1,\ldots,k})(p). \tag{2.10}$$

We can also consider $\alpha_1 \wedge \cdots \wedge \alpha_k$ to be an object acting on k-vector fields, i.e.

$$\begin{aligned}(\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_k)_p(v_1, \ldots, v_k) &:= \langle \alpha_1(p) \wedge \cdots \wedge \alpha_k(p), v_1(p) \wedge \cdots \wedge v_k(p) \rangle \\ &= \det(\{\alpha_i(v_j)\}_{i,j=1,\ldots,k})(p).\end{aligned}$$

Finally, we are ready to define differential forms on M:

Definition 2.14 (Differential Forms)**.** A differential form ω on M is an element of $\Gamma(\wedge(T^*M)) := \Omega(M)$.

$\Omega(M)$ is a $\mathbb{Z}$-graded, graded commutative algebra, which can be decomposed into homogeneous components: $\Omega^k(M) = \Gamma(\wedge(T^*M)^k)$. Now let's see some examples.

Example 2.1. *If $M = \mathbb{R}^3$, $\omega = 3 + dx^2 + dx^1 \wedge dx^3 - 5(x^2)^3 dx^1 \wedge dx^2 \wedge dx^3 \in \Omega(\mathbb{R}^3)$.*

Another important example, which we will see a lot is the following:

Example 2.2. *Let $f : M \to \mathbb{R}$ be a smooth function. Then $df \in \Gamma(T^*M) = \Omega^1(M)$ (i.e. a "one-form"). So, given a function f on our manifold, at each point we get a map that sends a vector in T_pM to a real number, because it gives a dual vector at every point. Thinking globally, you can see this as an object which takes in a vector field on $M(\in \Gamma(TM))$ and spits out another function on M.*

If we are mapping, instead, to another manifold N, so $f : M \to N$, there is a naturally induced map from differential forms on N back to M, which is called the *pullback* of f, and defined to be

Definition 2.15 (Pullback)**.** Let $f : M \to N$ be a smooth function, $\omega \in \Omega^k(N)$, and $v_1, \ldots, v_k$ be k vector fields on M. Then the **pullback** of f acting on ω, $f^*\omega \in \Omega^k(M)$ is defined to be:

$$(f^*\omega)_p(v_1, \ldots, v_k) = \omega_{f(p)}(df_p(v_1), \ldots, df_p(v_k)). \tag{2.11}$$

The pullback has the following properties: $(f \circ g)^*\omega = g^* f^* \omega$, $f^*(\alpha \wedge \beta) = f^*\alpha \wedge f^*\beta$. I will now define the *exterior derivative*, which will be our last important tool needed from differential topology before we move on to introducing symplectic topology. One can locally express a differential k-form $\omega \in \Omega^k(M)$ in terms of coordinates given by a chart (U, x) by $(x^{-1})^*\omega = \sum_{i_1 < \cdots < i_k} a_{i_1,\ldots,i_k}(x(p)) dx_{i_1} \wedge \cdots \wedge dx_{i_k}$. Now, on this affine space, we define the exterior derivative in affine space:

$$d((x^{-1})^*\omega) = \sum_{i_1 < \cdots < i_k, \ell} \frac{\partial a_{i_1,\ldots,i_k}}{\partial x^\ell}(x(p)) dx_\ell \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}. \tag{2.12}$$

All we have to do to get the **exterior derivative** is now map back to M:

$$d\omega = x^* d((x^{-1})^* \omega). \tag{2.13}$$

As you can see in Eq. 2.12, this operator takes a k form and creates a $k+1$ form. Applying d to a function ($\in \Omega^0(M)$) gives us a one form, and we recover the derivative we already know, df. Note that d is linear over $\mathbb{R}$. Here are some properties of d which are important:

1. $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^{|\alpha|}\alpha \wedge d\beta$, where $\alpha \in \Omega^{|\alpha|}$.
2. $d^2 = 0$.
3. $f^*(d\omega) = d(f^*\omega)$.

These properties follow from manipulations on affine space, given the definition of d I gave in Eq. 2.12. If an arbitrary form $\omega \in \Omega^k(M)$ is such that $d\omega = 0$, then we call ω a **closed differential form**. Furthermore, ω is **exact** if $\omega = d\alpha$, for some $\alpha \in \Omega^{k-1}(M)$.

2.1.2 Symplectic Manifolds

Hamiltonian dynamics commonly occur on special kind of manifolds called *Poisson manifolds*, but the dynamics are constrained to be on what are called *symplectic* submanifolds (see Subsection 2.2.1 for details) which have special structure given to them by a two-form $\omega \in \Omega^2(M)$, very much like $\mathbb{R}^3$ has some greater structure given to it by the existence of the inner product, which defines a metric. In what follows, I will define this special kind of manifold, namely a *symplectic manifold*, describe/prove some of its properties and then go on to talk about Poisson manifolds, which will finish out the section.

Since we are dealing with a two-form, which are naturally defined pointwise at each tangent space T_pM, it is convenient to first define some terminology for general vector spaces and then generalize this to a form acting on tangent spaces at each point.

Definition 2.16 (Symplectic Map)**.** Let V be an m-dimensional vector space over $\mathbb{R}$. A skew-symmetric, bilinear map $A : V \times V \to \mathbb{R}$ is **symplectic** if it is non-degenerate. These properties can be written out explicitly:

1. *(Bilinear)* It is linear for each slot: for $\alpha, \beta \in \mathbb{R}$ and $u, v, w \in V, A(\alpha v + \beta w, u) = \alpha A(v, u) + \beta A(w, u)$. And similarly for the other slot.
2. *(Skew-symmetric)* $A(v, w) = -A(w, v)$, for all $u, v \in V$.
3. *(Non-degenerate)* Fix $v \in V$. If $A(v, w) = 0$ for all $w \in V$, then $v = 0$.

One important consequence to note is that the map $\tilde{A} : V \to V^*$, defined by $\tilde{A}(v) = A(v, \cdot) \in V^*$, is an *isomorphism*. Another thing to note is that, if we have such a symplectic map A on a vector space V, $\dim V$ must be *even*.

Theorem 2.17. *Let V be a vector space and A a symplectic map on V. Then $\dim V$ is even.*

Proof. We are free to choose any basis for V, making A a matrix with real coefficients A_{ij}. The skew-symmetric condition tells us that (using Einstein summation convention)[4]

$$\begin{aligned} A_{ij}v^i w^j &= -A_{ij}w^i v^j \\ &= -A_{ji}v^i w^j \\ &\iff \\ A &= -A^t. \end{aligned}$$

So we have that A is equal to the negative of its transpose. Now, we can look at the determinant. I'd like to point three properties of the determinant before I go on: $\det(A) = \det(A^t)$, $\det(-A) = (-1)^{\dim V}\det(A)$, and that the determinant is independent of coordinate system. So, using these facts, as well as the result from skew-symmetry, we get:

$$\det(A) = \det(A^t) = \det(-A) = (-1)^{\dim V}\det(A).$$

If $\dim V$ is odd, then we get that a real number is equal to negative itself, telling us that $\det(A) = 0$, which is a contradiction since it is non-degenerate. Therefore, $\dim V$ must be even. □

Now, we can easily generalize this to 2-forms on manifolds. Let $\omega \in \Omega^2(M)$ be a 2-form on M. Then by definition, ω_p is a map on the vector space $T_pM \times T_pM$. This motivates the following definition:

Definition 2.18 (Symplectic Form)**.** Let M be a smooth m-manifold, and $\omega \in \Omega^2(M)$. We say that ω is **symplectic** if ω is closed and ω_p is a symplectic map for all $p \in M$.

Given a symplectic form $\omega \in \Omega^2(M)$, the pair (M, ω) is called a **symplectic manifold**. Let's see some examples of symplectic manifolds.

Example 2.3. *Let $M = \mathbb{R}^{2m}$ with coordinates $\{q_1, \ldots, q_m, p_1, \ldots, p_m\}$, and $\omega_0 = dq_i \wedge dp_i$ (again, implied sum). Then $(\mathbb{R}^{2m}, \omega_0)$ is a symplectic manifold.*

[4] Unless specified otherwise, I will always be using Einstein summation convention.

Example 2.4. *Let $M = S^2$ and $\omega \in \Omega^2(S^2)$ be defined on spherical coordinates $\{\theta, \phi\}$ as $\omega = \sin\theta d\theta \wedge d\phi$. Although the coordinate system does not cover all of S^2, we can define it the same way in another coordinate system in which ϕ is translated by some non-zero α. Furthermore, this form can be extended to the poles by changing to projection coordinates and then evaluating at $(x, y) = (0, 0)$. (M, ω) is also a symplectic manifold. In fact, if we think of the point (θ, ϕ) as some unit vector $\hat{x} \in S^2$ and $v, w \in T_{\hat{x}}S^2$, then the form ω evaluated at $\hat{x}$, v, and w is: $\omega_p(v, w) = \hat{x} \cdot (v \times w)$.*

Since we will be looking at the evolution of our space or point of interest in time, given by some flow g_t, say, then we are also interested in studying functions to and from symplectic manifolds. In simple differentiable manifolds, we cared about smooth, bijective functions with smooth inverses, i.e. diffeomorphisms. However, now we are not only interested in diffeomorphisms of M, but also in functions which somehow leave our symplectic form fixed. We call this a *symplectomorphism*:

Definition 2.19 (Symplectomorphism)**.** Let (M_1, ω_1) and (M_2, ω_2) be two symplectic manifolds. A **symplectomorphism** $\phi : M_1 \to M_2$ is a diffeomorphism such that $\phi^*\omega_2 = \omega_1$.

One very interesting result of studying classifications of symplectic manifolds under symplectomorphisms is the fact that *every symplectic manifold (M, ω) of dimension $2m$ is locally symplectomorphic to $(\mathbb{R}^{2m}, \omega_0)$, where $\omega_0 = dq_i \wedge dp_i$.* This is called the Darboux Theorem.

Theorem 2.20. *[Darboux] Let (M, ω) be a $2m$-dimensional symplectic manifold. Then, around any point $p \in M$, there exists a chart $(U, \phi = \{q_1, \ldots, q_m, p_1, \ldots, p_m\})$ containing p such that on U, $(\phi^{-1})^*\omega = \omega_0 = dq_i \wedge dp_i$.*

In a way, this somehow implies that locally there isn't anything interesting going on. However, this is good news for physicists because, as you will see, this means we can always make a suitable change of coordinates and recover Hamilton's equations. It is also good news for mathematicians because it means that, if one can prove some assertion locally on $(\mathbb{R}^{2m}, \omega_0)$ and show that it is invariant under symplectomorphisms, then it must hold for any symplectic manifold.

The non-degeneracy of ω is another useful property of a symplectic form because it allows us to associate a vector field $v \in \Gamma(TM)$ to a 1-form $\beta_v \in \Omega^1(M)$ via the isomorphism $\tilde{\omega} : \Gamma(TM) \to \Omega^1(M)$ defined to be $\tilde{\omega}(v) = \omega(v, \cdot) = \iota_v\omega$. Since this is an isomorphism, we can also go the other way and define a vector field, given a one form $\beta \in \Omega^1(M)$, by $\tilde{\omega}^{-1}(\beta) = v_\beta \in \Gamma(TM)$. Of particular interest to us is the special case of defining

vector fields given *exact* one forms, that is, $\beta = dH$ for $H : M \to \mathbb{R}$. This brings us to an important definition.

Definition 2.21 (Hamiltonian Vector Field)**.** Let (M, ω) be a symplectic manifold and let $H : M \to \mathbb{R}$ be a smooth function on M. A **Hamiltonian vector field** X_H with **Hamiltonian function** H is the vector field defined by

$$X_H = \tilde{\omega}^{-1}(dH). \tag{2.14}$$

This is equivalent to saying that $\iota_{X_H}\omega = dH$.

If we now suppose that either M is compact or that X_H is complete, using Theorem 2.11, we can integrate this vector field and get a one-parameter family of diffeomorphisms $g_t^H : M \to M$ such that at $t = 0$ we have the identity and the velocity of the flow at each point $p \in M$ and $t \in \mathbb{R}$ is $X_H(p)$. In fact, this flow is more than just a diffeomorphism, it's a *symplectomorphism*.

Theorem 2.22. *Let (M, ω) be a symplectic manifold and $H : M \to \mathbb{R}$ be our smooth function on M which has a corresponding flow, $g_t^H : M \to M$, given by X_H. Then g_t^H is a symplectomorphism, i.e.*

$$\left(g_t^H\right)^* \omega = \omega \quad \forall\, t \in \mathbb{R}. \tag{2.15}$$

Unfortunately, the proof requires a notion of Lie derivatives, so I will not go into this. For a proof consult either *Arnold* [1] or *da Silva* [9]. Now we start seeing some connection to dynamics! We have just generated a flow on M via a function $H \in C^\infty(M)$. **From now on, if all I say is $F \in C^\infty(M)$, then will be $F : M \to \mathbb{R}$. Otherwise I will make it clear, such as by saying $G : M \to N$.** Let's see a quick example of such a flow on a familiar object and with a one form we have already defined:

Example 2.5. *Recall Example 2.4, where we had (S^2, ω) as our symplectic manifold, where $\omega = \sin\theta d\theta \wedge d\phi$ on the spherical coordinates patch (which we will work on here). Let's look at the flow on the 2-sphere that is generated by the height function $H(\theta, \phi) = \cos\theta$ on this patch. Taking dH we get $dH = \frac{\partial H}{\partial \theta} d\theta + \frac{\partial H}{\partial \phi} d\phi = -\sin\theta d\theta$. Can we find a vector field X_H on S^2 such that $\omega(X_H, \cdot) = dH = -\sin\theta d\theta$? We want to solve the following equation: $(\sin\theta d\theta \wedge d\phi)(X_H, \cdot) = -\sin\theta d\theta$. Now, $\sin\theta d\theta \wedge d\phi$ can be written as $\sin\theta(d\theta \otimes d\phi - d\phi \otimes d\theta)$, where here the tensor product means the following: $(\alpha \otimes \beta)(v, w) = \alpha(v)\beta(w)$. Looking at this form, we can easily see that defining $X_H = \frac{\partial}{\partial \phi}$ will get us what we want:*

$$\sin\theta(d\theta \wedge d\phi)(\partial/\partial\phi, \cdot) = -\sin\theta d\theta = dH.$$

Therefore, H generates a flow along $\partial/\partial\phi$, which is one that goes around the vertical axis.

Note that in the previous example, the flow was one which *preserved* the height, so $H(\theta, \phi)$ did not change with the flow generated by $\partial/\partial\phi$. In fact, this is a general property of such flows, which allows us to go further into our tie with dynamics and show that a flow X_H generated by H will *preserve* H along the flow.

Theorem 2.23. *Let (M, ω) be a symplectic manifold and $H : M \to \mathbb{R}$ be our smooth function on M which has a corresponding flow, $g_t^H : M \to M$, given by X_H. Then*

$$\frac{d}{dt}\left(g_t^H\right)^* H = \frac{d}{dt}\left(H \circ g_t^H\right) = 0, \quad \forall\ t \in \mathbb{R}. \tag{2.16}$$

And so, for all $p \in M$,

$$H(p) = \left(\left(g_t^H\right)^* H\right)(p) = H\left(g_t^H(p)\right). \tag{2.17}$$

Proof. Note the following:

$$\begin{aligned}\frac{d}{dt}\left(H \circ g_t^H\right)(p) &= d\left(H \circ g^H\right)_{(p,t)}\left(\frac{\partial}{\partial t}\right) \\ &= (dH)_{\left(g_t^H(p)\right)} \circ \left(dg^H\right)_{(p,t)}\left(\frac{\partial}{\partial t}\right).\end{aligned}$$

Now, by definition,

$$\left(dg^H\right)_{(p,t)}\left(\frac{\partial}{\partial t}\right) = X_H\left(g_t^H(p)\right),$$

so this gives us:

$$\begin{aligned}\frac{d}{dt}\left(H \circ g_t^H\right)(p) &= (dH)_{\left(g_t^H(p)\right)} \circ \left(dg^H\right)_{(p,t)}\left(\frac{\partial}{\partial t}\right) \\ &= dH_{\left(g_t^H(p)\right)}\left(X_H\left(g_t^H(p)\right)\right) \\ &= \omega_{\left(g_t^H(p)\right)}\left(X_H\left(g_t^H(p)\right), X_H\left(g_t^H(p)\right)\right) = 0.\end{aligned}$$

On the last line I used Definition 2.21. Since we proved this for an arbitrary $t \in \mathbb{R}$ and $p \in M$, it must be true for all $t \in \mathbb{R}$ and $p \in M$, thereby proving the theorem. □

This will turn out to be our conservation of energy along a flow generated by a Hamiltonian[5]!

[5] Hamiltonian will be time-independent in everything we deal with. See *Arnold* [1] for a generalization for when this is not the case.

Next I will demonstrate the symplectic nature of the cotangent bundle to a manifold P[6], T^*P. We will be looking to define a symplectic form $\omega \in \Omega^2(T^*P)$ – that is, a 2-form *on* T^*P, which takes two elements of $T(T^*P)$ (the tangent bundle of the cotangent bundle) and spits out an element of $C^\infty(T^*P)$. Note the subtleties! I will define ω in a coordinate free way, but then use coordinates to show its familiar form. Recall that T^*P is a vector bundle over P, which means it has a projection $\pi : T^*P \to P$, which takes $q = (p, v) \in T^*P$ to $p \in P$, where $v \in T_p^*P$. Taking the differential of π at a point q, we get $d\pi_q : T_q(T^*P) \to T_pP$. We define the **tautological 1-form** $\alpha \in \Omega^1(T^*P)$ point wise by:

$$\alpha_q = \xi \circ d\pi_q, \tag{2.18}$$

for $q = (p, \xi) \in T^*P$ and $\xi \in T_p^*P$. What this one form essentially does point wise is take a tangent vector of T^*P, project it down to a tangent vector of P, then apply ξ to that vector to get a real number. See Figure 2.3 to get a clear picture and explanation. With this, we define the **canonical symplectic 2-form** ω on T^*P as

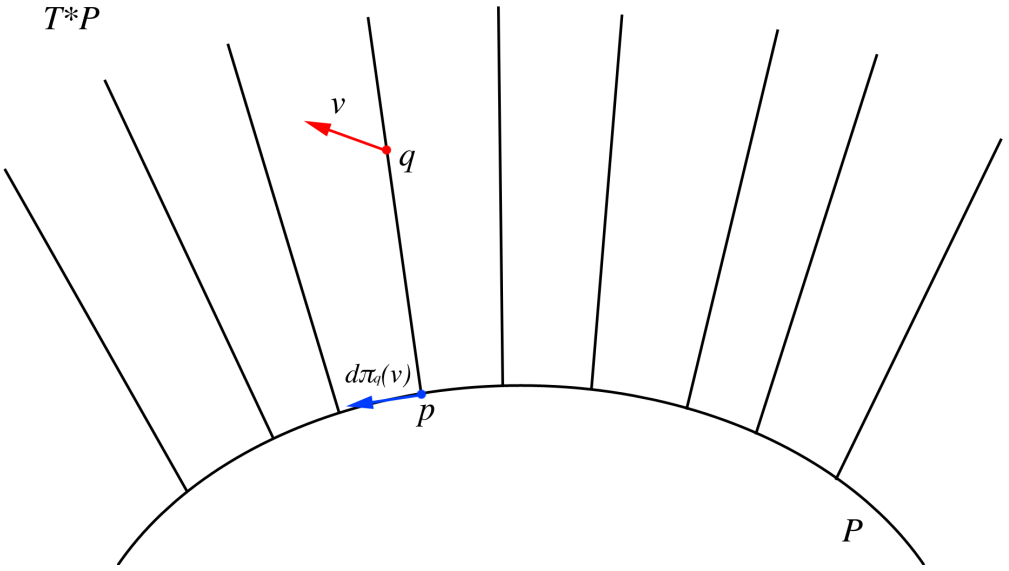


FIGURE 2.3: This sketch helps the reader understand what exactly the tautological 1-form α is doing at every point $q = (p, \xi) \in T^*P$. For $v \in T_q(T^*P)$, $\alpha_q(v)$ takes in v, projects it down to the corresponding tangent vector of P, and then applies $\xi \in T_p^*P$ to it to get a real number. Since this done for each $q \in T^*P$, α as a whole takes in an element of $\Gamma(T(T^*P))$, i.e. a vector field on T^*P, and spits out a function $C^\infty(T^*P)$.

$$\omega = -d\alpha. \tag{2.19}$$

Now that we have defined this 2-form without coordinates, let's look at what it looks like on a given coordinate patch of T^*P. Suppose we have a coordinate patch $(U, \phi =$

[6]I won't call it M to avoid confusion, because in previous examples M has been the symplectic manifold, whereas here T^*P is.

$\{x^1, \ldots, x^m\}$) on P around the point $p \in P$. On $T^*P|_U$, we have an induced basis $\{dx^1, \ldots, dx^m\}$, so a one-form $\xi \in T^*P|_U$ can be expanded to $\xi = \xi_i dx^i$ and so, in coordinates we have: $(p, \xi) \mapsto (x^1(p), \ldots, x^m(p), \xi_1, \ldots \xi_m) \in \mathbb{R}^{2n}$. Now, we can write $v \in T_q(T^*P)$ as $v = v^1\partial/\partial x^1 + \cdots + v^m\partial/\partial x^m + v^{m+1}\partial/\partial\xi^{m+1} + \cdots + v^{2m}\partial/\partial\xi^{2m}$. So, $d\pi_q(v) = v^1\partial/\partial x^1 + \cdots + v^m\partial/\partial x^m \in T_pP$. Then, $\xi(d\pi_q(v)) = \xi_i dx^i(v^1\partial/\partial x^1 + \cdots + v^m\partial/\partial x^m) = \xi_i v^i$, which tells us that, in coordinates,

$$(\phi^{-1})^*\alpha = \xi_i dx^i,$$

thereby making

$$(\phi^{-1})^*\omega = dx^i \wedge d\xi_i. \tag{2.20}$$

We have now shown that (T^*P, ω) is a symplectic manifold! We will use this fact in Subsection 2.2.1.

2.1.3 Poisson Manifolds

There is one more major structure which comes up in the study of Hamiltonian systems that is worth talking about. In fact, while it is easier to formulate theory behind Hamiltonian systems with symplectic manifolds, it turns out that "the natural scenery of almost all examples is a *Poisson manifold*" [2]. As we will see, a symplectic manifold is a special kind of Poisson manifold, so we are essentially generalizing the space on which a Hamiltonian system may be defined. Furthermore, we will see that a Poisson manifold is foliated by symplectic submanifolds on which the dynamics are constrained to lie on. Let's begin with the cold, hard definition of a Poisson manifold, and then connect it to what we have seen, and then shed some light on the implications of this generalization.

One quick thing to note before we begin: $C^\infty(M)$ has an algebra structure – it is a vector space over the field of real numbers, and has a product which is simply multiplying the two functions point wise: for $F, G \in C^\infty(M)$, $(F \cdot G)(p) = F(p)G(p) \in C^\infty(M)$.

Definition 2.24 (Poisson Manifold)**.** A **Poisson manifold** is a manifold M equipped with a **Poisson structure**. A Poisson structure is a bracket $\{\cdot, \cdot\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$ having the following properties:

1. *(Skew-symmetry)* Let $F, G \in C^\infty(M)$, then $\{F, G\} = -\{G, F\}$.
2. *(Bilinear)* Let $\alpha, \beta \in \mathbb{R}$, then $F, G, H \in C^\infty(M)$, then $\{\alpha F + \beta G, H\} = \alpha\{F, H\} + \beta\{G, H\}$.

3. *(Jacobi Identity)* Let $F, G, H \in C^\infty(M)$. Then

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0. \tag{2.21}$$

4. *(Leibniz Rule)* Let $F, G, H \in C^\infty(M)$. Then

$$\{F, GH\} = \{F, G\}H + G\{F, H\}. \tag{2.22}$$

The Leibniz rule allows us to associate the function F to a vector field X_F. How so? Well, one can define the action of a vector field $v \in \Gamma(TM)$ on a function $F \in C^\infty(M)$. We have $v : C^\infty(M) \to C^\infty(M)$ defined as the following: $v \cdot F = dF(v) \in C^\infty(M)$. So, applying a vector field to a function F gives us another smooth function which tells us, at each point $p \in M$, how F is changing in the direction of $v(p)$. One can go the other way and *associate a vector field to a function* $D : C^\infty(M) \to C^\infty(M)$ (look familiar?), but it must have two special properties – it must be a linear, and what's called a **derivation**, meaning that $D(FG) = FD(G) + D(F)G$ (intuitively, derivatives have these property, so it kind of makes sense). If we now look at the Leibniz Rule, we see that $-\{F, \cdot\}$[7] is a derivation, which means it defines a vector field, which we will call X_F, like before. Given this vector field, we can define a flow $g_t^H : M \to M$, as well. Furthermore, this implies that, for $H, F \in C^\infty(M)$,

$$X_H \cdot F = \{F, H\} = dF(X_H). \tag{2.23}$$

Thus $\{F, H\}$ is telling us how F changes along the flow generated by H. In fact, we come to the same result as before when we associated a function with a vector field using the symplectic form:

Theorem 2.25. *Let M be a Poisson manifold. The flow generated by the function H, found by integrating the vector field $X_H = -\{H, \cdot\}$, preserves the function H.*

The proof is very similar to that of Theorem 2.23.

Proof. We want to show that $\frac{d}{dt}(H \circ g_t^H) = 0$. In the proof of Theorem 2.23, we showed that this expression simplifies to $dH(X_H) = 0$. But by Eq. 2.23, this is equal to $\{H, H\} = 0$ due to skew-symmetry. Therefore, $dH(X_H) = 0$. □

Let's see an example of a Poisson manifold.

Example 2.6 (Lie-Poisson on $\mathbb{R}^3$)**.** *Let $P = \mathbb{R}^3$, which has the usual cross product $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$, and $M = (\mathbb{R}^3)^*$. Call the duality isomorphism $D : \mathbb{R}^3 \to (\mathbb{R}^3)^*$.*

[7] The sign is convention, but it is important.

One can define a Poisson structure $\{\cdot,\cdot\} : C^\infty((\mathbb{R}^3)^*) \times C^\infty((\mathbb{R}^3)^*) \to C^\infty((\mathbb{R}^3)^*)$ *on* $(\mathbb{R}^3)^*$ *by using this cross product and the duality pairing* $\langle\cdot,\cdot\rangle$*. Consider the coordinate one-forms,* $dx, dy, dz \in (\mathbb{R}^3)^*$*. Then we can define a Poisson bracket by:*

$$\{dx, dy\}(\cdot) = \langle \cdot, D^{-1}(dx) \times D^{-1}(dy)\rangle. \tag{2.24}$$

For those of you who know about Lie algebras, what we are doing is taking the Lie algebra structure of $\mathbb{R}^3$ *and using it to give a Poisson structure on* $(\mathbb{R}^3)^*$*. Also note that this manifold is of* **odd** *dimension! Recall that this was not possible for symplectic manifolds.*

This Poisson structure has given us the same sense of dynamics (that is, a flow resulting from associating a function to a vector field) as that of a symplectic manifold. So is there a connection? Yes! Suppose that the bracket on a Poisson manifold M is non-degenerate (we never said it had to be!), then we can define a symplectic structure on M, as well, by:

$$\{F, G\} = \omega(X_F, X_G) = dF(X_G). \tag{2.25}$$

It is not obvious that the conditions on $\{\cdot,\cdot\}$ give us the conditions that are necessary for our defined ω to be symplectic. We have bilinearity, skew-symmetry, and non-degeneracy automatically, but we also need ω to be closed! This will be given to us by the fact that our bracket satisfies the Jacobi identity. In the section on symplectic manifolds, we saw that Darboux's theorem tells us that there is a canonical coordinate patch $(U, \phi = \{q_1, \ldots, q_m, p_1, \ldots, p_m\})$ around any point of any symplectic manifold (M, ω) such that $\left(\phi^{-1}\right)^* \omega = dq_i \wedge dp_i = \omega_0$. So, any (M, ω) is locally symplectomorphic to $(\mathbb{R}^{2m}, \omega_0)$. There is a corresponding "canonical bracket" for ω_0.

Example 2.7. *For the canonical coordinate system of a symplectic manifold,* $(\mathbb{R}^{2m}, \omega_0)$*, we can derive the "canonical bracket" by applying Eq. 2.25. Let* $F, G \in C^\infty(\mathbb{R}^{2m})$ *be arbitrary functions on* $\mathbb{R}^{2m}$*. Using the canonical symplectic form* ω_0*, we can find out what their corresponding Hamiltonian vector fields are. We can express* X_F *in terms of basis vectors* $\frac{\partial}{\partial q_i}, \frac{\partial}{\partial p_i}$*:* $X_F = f^1 \frac{\partial}{\partial q_1} + \cdots f^m \frac{\partial}{\partial q_m} + f^{m+1} \frac{\partial}{\partial p_1} + \cdots + f^{2m} \frac{\partial}{\partial p_m}$*. Let's begin by finding out what the expression for each* f^i *and* f^{i+m} *is.*

$$\begin{aligned}
dF &= \omega_0(X_F, \cdot) \\
\frac{\partial F}{\partial q_i} dq_i + \frac{\partial F}{\partial p_i} dp_i &= (dq_k \wedge dp_k)\left(f^i \frac{\partial}{\partial q_i} + f^{i+m} \frac{\partial}{\partial p_i}, \cdot\right) \\
&= (dq_1 \wedge dp_1)\left(f^i \frac{\partial}{\partial q_i} + f^{i+m} \frac{\partial}{\partial p_i}, \cdot\right) + \cdots + \\
&\qquad (dq_m \wedge dp_m)\left(f^i \frac{\partial}{\partial q_i} + f^{i+m} \frac{\partial}{\partial p_i}, \cdot\right) \\
&= f^i dp_1 - f^{1+m} dq_1 + \cdots + f^m dp_m - f^{2m} dq_m \\
&= -f^{i+m} dq_i + f^i dp_i.
\end{aligned}$$

So, we see that

$$X_F = \frac{\partial F}{\partial p_i}\frac{\partial}{\partial q_i} - \frac{\partial F}{\partial q_i}\frac{\partial}{\partial p_i}. \tag{2.26}$$

Now we can calculate the bracket by using $\{G, F\} = dG(X_F)$*:*

$$\begin{aligned}
\{F, G\} &= -dG(X_F) \\
&= -dG\left(\frac{\partial F}{\partial p_i}\frac{\partial}{\partial q_i} - \frac{\partial F}{\partial q_i}\frac{\partial}{\partial p_i}\right) \\
&= -\left(\frac{\partial F}{\partial p_i}dG\left(\frac{\partial}{\partial q_i}\right) - \frac{\partial F}{\partial q_i}dG\left(\frac{\partial}{\partial p_i}\right)\right) \\
&= \frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}.
\end{aligned}$$

So we end up with the **canonical Poisson bracket**

$$\{F, G\} = \frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}. \tag{2.27}$$

To simplify this expression, some people write the following: $\vec{z} = (q_1, \ldots, q_m, p_1, \ldots, p_m)$ *(so we will sum from 1 to* $2m$ *now), and define:*

$$\{F, G\} = \frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i} = \frac{\partial F}{\partial z_i}J_c^{ij}\frac{\partial G}{\partial z_j}, \tag{2.28}$$

where $\mathbf{J}_c$ *is the (canonical) "cosymplectic form" defined by:*

$$\mathbf{J}_c = \begin{pmatrix} 0_m & I_m \\ -I_m & 0_m \end{pmatrix}, \tag{2.29}$$

and I_m *is the* $m \times m$ *identity matrix.*

So for certain Poisson manifolds, we also have symplectic manifolds, as we just saw – but what about the case where $\{\cdot, \cdot\}$ is degenerate? That is, there exist non-zero functions whose bracket with *anything* is *always* zero. These functions are of interest and so they deserve their own name and definition!

Definition 2.26 (Casimir Function)**.** Let M be a smooth Poisson manifold. A **Casimir function** (or just **Casimir**) of the Poisson bracket is a function $C \in C^\infty(M)$ such that

$$\{C, F\} = 0 \quad \forall\ F \in C^\infty(M). \tag{2.30}$$

Note that, due to this property, we also have that $X_F \cdot C = \{F, C\} = dC(X_F) = 0$. Therefore, these Casimirs are *preserved by any flow!* So, when we get to actual dynamics, they will become constants of motion! This implies that, along any flow, *we are restricted to submanifolds on which* C *(there could be more than one) is constant*, i.e. level sets

of the Casimirs of our bracket. This creates a foliation of the Poisson manifold by such submanifolds. Furthermore, it turns out that these submanifolds themselves are symplectic (and so of even dimension). Finally, there is a theorem (called Lie-Darboux by some) which is very similar to that of Theorem 2.20 which tells us that around any point in a Poisson manifold M with degenerate bracket, there is a chart $(V, \psi = \{q_1, \ldots, q_k, p_1, \ldots, p_k, u_1, \ldots, u_{m-2k}\})$ such that $\{F, G\}$ becomes the canonical Poisson bracket when restricted to the first $2k$ coordinates q_i, p_i [24], or, put another way, the cosymplectic form defined in Eq. 2.31 becomes:

$$\mathbf{J} = \begin{pmatrix} 0_k & I_k & 0_{m-2k} \\ -I_k & 0_k & 0_{m-2k} \\ 0_{m-2k} & 0_{m-2k} & 0_{m-2k} \end{pmatrix}. \tag{2.31}$$

2.1.4 Dirac Constraint Theory

We just saw in the last section that, along any flow generated by a function H and degenerate bracket $\{\cdot, \cdot\}$, the Casimir functions C^α ($\alpha \in \mathcal{A}$ being an index) remain constant. Therefore, given some Poisson manifold with a degenerate Poisson bracket $\{\cdot, \cdot\}$, there exists some set of Casimir functions that result from the definition of the bracket. Suppose now that you wanted to *require* that some non-Casimir function $F \in C^\infty(M)$ be constant along any flow. You can't just set F to zero and assume that F will remain zero for all time ($dF(X_H)$ might be non-zero at some point along the flow), so what can you do? One would have to *modify* the bracket $\{\cdot, \cdot\}$ so that F becomes a Casimir. This is what Dirac constraint theory is about – given some (even) set of functions $\{C^\alpha\}_\alpha$ that you want to be preserved by the flow, there is a systematic way of altering the bracket to create a new bracket which restricts the flow to submanifolds defined by $\{C^\alpha = 0\}_\alpha$, so long as you start on those manifolds. Geometrically, this process is intuitively what one would think: suppose we have a Poisson manifold M and a function H which generates a vector field X_H which in turn generates a flow g_t^H. We start at a point $p \in S \subset M$, where S is a submanifold on which $C^\alpha = 0$ for $\alpha \in \mathcal{A}$. If we want our flow to keep us on S, then at each point on M you have to alter the vector field X_H so as to make it tangent to S (otherwise you would be "lifting" the flow off of the submanifold).

We can begin by drawing from an analogy in usual inner product vector spaces (which we *don't* necessarily have in the more general situation, so this is just an *analogy!*) using some basic knowledge from linear algebra. Suppose we have a vector space V of dimension m and a subspace $W \subset V$ of dimension n defined by $W = \{\vec{x} \in V \mid \vec{F}^\alpha \cdot \vec{x} = 0, \alpha = 1, \ldots, m-n\}$ where $\vec{F}^\alpha \cdot \vec{F}^\beta = \delta_{\alpha\beta}$. Now let $\vec{v} \in V$ – we want to project $\vec{v}$ on W,

so what we do is simply take away the parts of $\vec{v}$ that are orthogonal to this subspace, namely $(\vec{v}\cdot\vec{F}^\alpha)\vec{F}^\alpha$. Therefore we define the projected $\vec{v}$ by

$$\vec{v}_\parallel = \vec{v} - \sum_{\alpha=1}^{m-n}(\vec{v}\cdot\vec{F}^\alpha)\vec{F}^\alpha. \tag{2.32}$$

You can easily check that $\vec{v}_\parallel \in W$ by dotting it with each $\vec{F}^\alpha$ and seeing that it is equal to zero. We want to translate this idea and do something similar for the vector field X_H, which defines a flow that we want to keep on $S \subset M$. Note now that we have no notion of inner product on M, so we are forced to work with the Poisson bracket and differentials of our generating functions. However, even though we have no clear sense of projection in terms of an inner product, geometrically we are essentially doing the same thing and in fact the final equation will resemble Eq. 2.32. The full details are a bit involved; for a full account consult Kurt Sundermeyer's book *Constrained Dynamics* [31].

We want to restrict our vector field onto the tangent space of S, which is $T_pS = \cap_\alpha \ker(dC^\alpha)$. Furthermore, we will not be simply looking at X_H, but we will be always looking at how X_H *acts* on a function $G \in C^\infty(M)$, say, which will tell us how G changes in the direction of the flow. So, $X_H \cdot G = \{H, G\} = dG(X_H) = -\{G, H\} = -dH(X_G)$. Therefore, what we can do is split up the differential of our generating function H into one whose image lies on $T_pS = \cap_\alpha \ker(dC^\alpha)$ and the other which lies on $\bigoplus_\alpha \mathrm{im}(dC^\alpha)$. So, assuming each dC^α is linearly independent (which is reasonable according sources) and that there are only r of them, we define $dH = d\tilde{H} + \sum_{\alpha=1}^r v^\alpha dC^\alpha$, where $v^\alpha \in C^\infty(M)$ act like Lagrange multipliers. We are defining $\tilde{H}$ to be the generating function whose flow lies on S. If we now apply a vector field X_G generated by $G \in C^\infty(M)$, we get:

$$\begin{aligned} dH(X_G) &= d\tilde{H}(X_G) + \sum_{\alpha=1}^r v^\alpha dC^\alpha(X_G), \\ \{H, G\} &= \{\tilde{H}, G\} + \sum_{\alpha=1}^r v^\alpha \{C^\alpha, G\}, \\ \{\tilde{H}, G\} &= \{H, G\} - \sum_{\alpha=1}^r v^\alpha \{C^\alpha, G\}. \end{aligned}$$

By definition, our the vector field $X_{\tilde{H}}$ is on the tangent space of S, so any of the constraint functions C^α will not change in this direction. This tells us that $\{\tilde{H}, C^\alpha\} = 0$

for all $\alpha \in \mathcal{A} = \{1, \ldots, r\}$. This allows us to solve for the v^α's:

$$\begin{aligned} \{H, C^\beta\} &= \sum_{\alpha=1}^{r} v^\alpha \{C^\alpha, C^\beta\} := \sum_{\alpha=1}^{r} v^\alpha \omega^{\alpha\beta} \\ &\iff \\ v^\alpha &= \sum_{\beta=1}^{r} \{H, C^\beta\} \left(\omega^{\beta\alpha}\right)^{-1}. \end{aligned} \tag{2.33}$$

where $\omega^{\alpha\beta} = \{C^\alpha, C^\beta\}$. The existence of this inverse was proven by Dirac, but it is worth commenting that I am brushing quite a bit of details under the rug by not discussing the rigorous derivation of everything, including what are called "primary constraints" and "secondary constraints" – for a detailed discussion see [31]. An important thing to note is that, since $\{C^\alpha, C^\beta\} = -\{C^\beta, C^\alpha\}$, $\omega^{\alpha\beta}$ is only invertible if there are an *even* number of constraint functions! Practically speaking, it is a common thing for one to start with a single constraint, but have to derive a second one so as to get an even number of constraints. New constraints can be found by looking at $\{H, C\}$, if C is the single constrain which you are starting with. If $\{H, C\} \neq 0$, then one can define $C^2 = \{H, C\}$, for example. Using Eq. 2.33, we now have that

$$\{\tilde{H}, G\} = \{H, G\} - \sum_{\alpha,\beta=1}^{r} \{H, C^\beta\} \left(\omega^{\beta\alpha}\right)^{-1} \{C^\alpha, G\}. \tag{2.34}$$

The left hand side will give us dynamics (generated by $\tilde{H}$) which are constrained to be on our submanifold of interest S, by construction. Notice, however that the right hand side is defined completely in terms of H, C^α's, and the old bracket. So all we need in order to get the correct dynamics is not some new function $\tilde{H}$, but rather a new *bracket*, defined in terms of H, the old bracket, and the functions C^α which we want to make Casimirs. This motivates us to define what we will call the **Dirac bracket**:

Definition 2.27 (Dirac Bracket)**.** Let M be a Poisson manifold with bracket $\{\cdot, \cdot\}$, $F, G \in C^\infty(M)$, and $\{C^\mu\}_1^{2k} \in C^\infty(M)$ the even-numbered set of constraints that we are trying to make Casimirs. Then the **Dirac bracket** for our Poisson manifold, denoted by $\{\cdot, \cdot\}_D$, is defined to be:

$$\{F, G\}_D = \{F, G\} - \sum_{\mu,\nu=1}^{2k} \{F, C^\mu\} \left(\omega^{\mu\nu}\right)^{-1} \{C^\nu, G\}. \tag{2.35}$$

I will not show any examples yet, simply because the Hamiltonian formulation of the Riemann ellipsoids is one such example, and you will get to see the Dirac bracket in action then.

We are finally done with the required mathematical background and will move on to tie all of this to physics. As you can probably tell, the step towards physics is not too large since we have already been dealing with ideas of dynamics and flows. In the next section I will make it explicit how Hamiltonian dynamics come into play, using all of the tools we just built up.

2.2 Hamiltonian Theory

2.2.1 Hamiltonian Dynamics

In the previous section we studied the properties of local flows that were generated by arbitrary smooth functions on our manifold M. In real-world Hamiltonian systems you specify *one* which is called (you guessed it) the *Hamiltonian* H. As you will see, in many physical cases the Hamiltonian will be the energy, for example.

Definition 2.28 (Hamiltonian System)**.** Let $(M, \{\cdot,\cdot\})$ be a smooth, compact Poisson manifold and $H : M \to \mathbb{R}$ be a smooth function on M. Then the triplet $(M, \{\cdot,\cdot\}, H)$ is **Hamiltonian system** with **Hamiltonian function** H. This Hamiltonian creates a vector field $X_H = -\{H, \cdot\}$ which in turn creates a one-parameter flow on an open set $U \subset M$, $g_t^H : U \to U$. We call the parameter t *time* and, for any $F \in C^\infty(M)$,

$$X_H \cdot F = dF(X_H) =: \dot{F}, \tag{2.36}$$

is "how F changes in time." Finally, the *equation of motion* for some function $F \in C^\infty(M)$ is:

$$\dot{F} = \{F, H\}. \tag{2.37}$$

An important thing to note is that **I will not be dealing with time-dependent Hamiltonians** in this thesis, but it does happen in some realistic situations. I recommend the reader to look at *Arnold* [1] for details on that. Notice that if $\{\cdot,\cdot\}$ is non-degenerate, then we get a triplet (M, ω, H), which also defines a Hamiltonian system, but a more specific case.

Theorem 2.29. *Let $(M, \{\cdot,\cdot\}, H)$ be a Hamiltonian system and $F \in C^\infty(M)$ be a smooth, real-valued function on M. F remains constant along the flow if and only If $\{F, H\} = 0$. We say that F does not change in time, or is a* **constant of motion**.

Proof. We want to show that $\frac{d}{dt}(F \circ g_t^H) = 0$. By following the proof of Theorem 2.23, one can see that this expression simplifies to $dF(X_H) = 0$. But by Eq. 2.23, this is equal to $\{F, H\}$, which is equal to 0 by assumption. Therefore, $\dot{F} = dF(X_H) = 0$. □

So, any function whose bracket with H is zero does not change in time. Therefore, we see that automatically any Casimir of the bracket is a constant of motion. There is an important subtlety to note here – Casimirs are constants of motion *regardless of the Hamiltonian*, but there may also be a set of functions $\{F^i\}_i$ such that $\{H, F^i\} = 0$ for all $i = 1, \dots k$ but not for a function other than H! One set of constants of motion (Casimirs) is due to the Poisson bracket, and the other set of constants of motion is due to the Hamiltonian itself. Both of course have physical reasons – for example, in the problem of rigid body rotation with a fixed point, the magnitude of the gravitational vector (which moves around in the frame of reference of the spinning body) is a Casimir of the bracket [2]. However, it makes total physical sense for that to be true, as the gravitational force is only changing direction in that frame. On the other hand, its angular momentum is a conserved quantity, but this is due to a symmetry in the Hamiltonian (see: Noether's Theorem).

Suppose we are working with a Hamiltonian system which has a non-degenerate Poisson bracket (i.e. a symplectic manifold), then Theorem 2.20 and Example 2.7 tell us that there exists a diffeomorphism which maps $(M, \{\cdot,\cdot\}, H)$ to $(\mathbb{R}^{2m}, \{\cdot,\cdot\}, H)$ with coordinates $\vec{z} = (q_1, \dots, q_m, p_1, \dots, p_m)$ and $\{\cdot,\cdot\} = \frac{\partial}{\partial z_i} J_c^{ij} \frac{\partial}{\partial z_j}$. Plugging in the coordinate functions q_i, H and p_i, H separately into Eq. 2.37 using this coordinate system, we get the famous **Hamilton's equations:**

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \tag{2.38}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{2.39}$$

Solving these equations for $\vec{z}(t)$ gives us the integral curves of the flow generated by H. But what exactly does this mean physically? So far I have claimed to be talking about "physics," but it doesn't seem I've introduced anything too physical, other than "choosing" a Hamiltonian function H and claiming that this gives us real-world dynamics. But how does one arrive at all of this from a purely physical perspective? Like in most classes, one begins with Lagrangian mechanics!

We are interesting in finding the motion of particle(s). We begin by specifying a space (manifold Q) on which our particles are confined – our **configuration manifold (or space)**. One typically does this by either parametrizing the space which you a priori know will be where your particle lives (say, if you are looking at a particle on a ring or on S^2), or after given some sort of constraint function, $f : Q \to \mathbb{R}$, $f = 0$ defining a level set on which your particle must reside. Either way, if one hopes to do any physics, it is essential that one is able to put a coordinate system on your configuration space Q. Let's say that our configuration space Q is m-dimensional and has coordinates

$\{q_1, \ldots, q_m\}$. Once you define your configuration space, you must appeal to the laws of physics which have been observed and empirically derived. These equations will give us a flow $g_t : Q \to Q = g : Q \times \mathbb{R} \to Q$ of the space, or, more physically, a trajectory for the particle, denoted by $\vec{q}(t)$[8]. This flow will have a velocity

$$\dot{q} = \dot{\vec{q}} = \frac{d}{dt} q(t) = dg_{(q_0,t)} \left(\frac{\partial}{\partial t} \right). \tag{2.40}$$

One very powerful idea of classical mechanics which results in equations of motion is that of an *action principle*. The most famous of which is called Hamilton's principle, which states the following:

Theorem 2.30 (Hamilton's Principle)**.** *Let Q be a configuration space with coordinates $q(t) = \vec{q}(t) = \{q_1(t), \ldots, q_m(t)\}$, $T : TQ \to \mathbb{R}$ kinetic energy, and $V : TQ \to \mathbb{R}$ the potential energy. Define the* **Lagrangian** *$L : TQ \to \mathbb{R}$ by $L = T - V$. The motion of a mechanical system $\vec{q}(t)$ are one-parameter curves on Q which are extremals of the* **action functional***:*

$$S[q] = \int_{t_0}^{t_1} L(q, \dot{q}) dt, \tag{2.41}$$

with fixed $q(t_0)$ and $q(t_1)$ as boundary conditions for the variation.

Note that the Lagrangian is not an explicit function of time (but q and $\dot{q}$ are). This is *not* the general case, but I will only deal with this situation for the sake of time. Again, for a discussion with full generality consult *Arnold* [1]. Finding the extremal of the action functional requires one to take a *functional derivative* of $S[q]$, which is defined as follows [24]:

Definition 2.31 (Functional Derivative)**.** The ith component of the **functional derivative** of $S[q] = \int_{t_0}^{t_1} L(q, \dot{q}) dt$, called $\delta S / \delta q_i$, is

$$\frac{\delta S}{\delta q_i} = \frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} + \frac{d^2}{dt^2} \frac{\partial L}{\partial \ddot{q}_i} - \cdots \tag{2.42}$$

Finding the extremal of $S[q]$ means we solve for $\delta S / \delta q_i = 0$, for each $i = 1, \ldots m$. In our case, this results in the famous **Euler-Lagrange equations**:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0 \tag{2.43}$$

By writing L in terms of coordinates of TQ we recover the equations of motion.

Example 2.8. *In a lot of examples, $T = \frac{1}{2} m (\dot{q})^2$ and V is a function of only the position q. These types of systems are called "natural systems." By sticking these definitions into*

[8] I sometimes choose to leave out the vector over the q for neatness. The key to tell whether or not I'm talking about the vector $\vec{q}$ or its components is to see whether or not I have placed an index.

the Euler-Lagrange equation, we get:

$$m\ddot{q}_i = -\frac{\partial V}{\partial q_i}. \tag{2.44}$$

This is Newton's second law!

To get to Hamilton's equations, we must do a *Legendre transformation*, which is a trick (seen a lot of times in thermodynamics) to change the functional dependence of a function. We begin by defining the *canonical momentum*:

Definition 2.32 (Canonical Momentum)**.** Given a Lagrangian system with Lagrangian $L(q, \dot{q})$, the **canonical momentum** of the system is

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{2.45}$$

In fact, written without referring to coordinates, one can view $p(q)$ as a *dual vector* on the tangent space T_qQ, by:

$$p(q) = dL_q, \tag{2.46}$$

and so in coordinates of T_q^*Q, $p(q) = p_i d\dot{q}_i$ and $dL_q = (\partial L/\partial \dot{q})|_q d\dot{q}_i$, which returns us to Eq. 2.45. Since this definition is true for all $q \in Q$, then we can extend p to be a *section of the cotangent bundle* T^*Q, i.e. $p \in \Omega^1(Q)$. We call T^*Q the **phase space** of our system, and it has coordinates $\{q_1, \ldots, q_m, p_1, \ldots, p_m\}$. A point in phase space, (q, p), tells us the *state* of the system, that is, all the information that is needed to solve for the dynamics $q(t)$ (along with initial conditions, of course).

Example 2.9. *In the case where* $T = \frac{1}{2}m\left(\dot{q}\right)^2$ *and* V *is a function of only the position* q*, our canonical momentum is*

$$\begin{aligned} p_i &= \frac{\partial L}{\partial \dot{q}_i} \\ &= \frac{1}{2}m\frac{\partial \left(\dot{q}\right)^2}{\partial \dot{q}_i} \\ &= m\dot{q}_i. \end{aligned}$$

One instantly recognizes this as the usual definition of momentum that many people learn in high school or beginning physics.

Example 2.10. *However, what we just saw in the previous example doesn't always happen. Take a system with* $T = \frac{1}{2}m\left(\dot{q}\right)^2$ *and* $V = eW(q) - \frac{e}{c}\dot{q}_i A_i$*, where* e *is the electric charge of the particle (electron),* A_i *is the magnetic vector potential, and* W *is*

the electric potential. In this case, the canonical momentum is

$$\begin{aligned} p_i &= \frac{\partial L}{\partial \dot{q}_i} \\ &= \frac{1}{2}m\frac{\partial \left(\dot{q}\right)^2}{\partial \dot{q}_i} + \frac{e}{c}\frac{\partial(\dot{q}_j A_j)}{\partial \dot{q}_i} \\ &= m\dot{q}_i + \frac{e}{c}A_i. \end{aligned}$$

With this definition of momentum, we now define the (physical) Hamiltonian function, which will generate Hamilton's equations, as we will see.

Definition 2.33 ((Physical) Hamiltonian)**.** Given a Lagrangian function $L : TQ \to \mathbb{R}$, the **(physical) Hamiltonian** of the system, $H : T^*Q \to \mathbb{R}$, is defined to be

$$H(q,p) = p_i\dot{q}_i(p) - L(q,\dot{q}). \tag{2.47}$$

In most "natural" situations, this Hamiltonian will be the total energy. Note that I have written H in terms of (q,p) (coordinates of T^*Q, the phase space), which means I have defined $\dot{q}$ in terms of p (as implied in the definition). This is *not* always doable, in fact, and relies on the invertibility of Eq. 2.45, which requires that L be what is called a *convex* function [24]. When this Legendre map is non-invertible, one ends up with what are called *dynamical constraints*, which are constraints on the *phase space* which is dynamically accessible (which is different than constraints we have seen above, like $f = 0$, which simply determined the manifold Q) [31]. These constraints might require something about the properties of the velocity, for example, and force the phase space flow to lie on a submanifold of T^*Q. As we have seen before, Dirac brackets allow us to define constrained dynamics, so we can use Dirac constraint theory to tackle this problem. Now that we have a Hamiltonian, let's derive Hamilton's equations. We begin by taking the differential to both sides of Eq. 2.47 and expanding out in terms of the coordinates of T^*Q.

$$\begin{aligned} dH(q,p) &= d(p_i\dot{q}_i(p) - L(q,\dot{q}(p))) \\ \frac{\partial H}{\partial q_j}dq_j + \frac{\partial H}{\partial p_j}dp_j &= p_i\frac{\partial \dot{q}_i(p)}{\partial p_j}dp_j + \dot{q}_j dp_j - \frac{\partial L}{\partial q_j}dq_j - \frac{\partial L}{\partial \dot{q}_i}\frac{\partial \dot{q}_i}{\partial p_j}dp_j, \\ \frac{\partial H}{\partial q_j}dq_j + \frac{\partial H}{\partial p_j}dp_j &= \left(p_i - \frac{\partial L}{\partial \dot{q}_i}\right)\frac{\partial \dot{q}_i(p)}{\partial p_j}dp_j + \dot{q}_j dp_j - \frac{\partial L}{\partial q_j}dq_j, \\ \frac{\partial H}{\partial q_j}dq_j + \frac{\partial H}{\partial p_j}dp_j &= -\frac{\partial L}{\partial q_j}dq_j + \dot{q}_j dp_j, \end{aligned}$$

where I used Eq. 2.45 on the last line. This tells us that

$$\begin{aligned} -\frac{\partial L}{\partial q_j} &= \frac{\partial H}{\partial q_j}, \\ \dot{q}_j &= \frac{\partial H}{\partial p_j}. \end{aligned}$$

Finally, using the Euler-Lagrange equation, we can write the final set of equations, which are in fact Hamilton's equations:

$$\dot{q}_i = \frac{\partial H}{\partial p_j}, \tag{2.48}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_j}. \tag{2.49}$$

This ties us back to where we were before we started looking into the real physics, connecting the physics to the natural symplectic structure of T^*Q with its canonical symplectic form (seen in Eq. 2.20) and its resulting canonical Poisson bracket, which gives Hamilton's equations. Since Darboux's theorem guarantee's that we can always transform our coordinates so as to make the symplectic form canonical, this is as general as one needs to be. Of course, in many situations, though, one does not have a clear way to transform to canonical coordinates and must then resort to not putting the equations of motion in the form of Hamilton's equations, but rather must use a physical Hamiltonian (usually simply the energy, if the system is "natural") and define a Poisson bracket for the system such that using Eq. 2.37 gives one the right equations of motion. It is important to note that there are many cases, called **non-canonical Hamiltonian systems** when this bracket $\{\cdot,\cdot\}$ which one defines depends on the point in phase space on which one evaluates it (not like the canonical Poisson bracket, which is independent of phase space location), and of course there are certain situations in which $\{\cdot,\cdot\}$ is degenerate and we don't have a symplectic manifold. Thankfully this is okay because none of the properties we derived depended on this canonical coordinate system except for Hamilton's equations, and in the case of a degenerate bracket, the dynamics are restricted to happen on a symplectic submanifold of our phase space, which then recovers the nice properties we have and will define, when restricted to that submanifold.

Having satisfied the thirst of the physicists, we go back to discussing Hamiltonians in the abstract context of the triplet $(M, \{\cdot,\cdot\}, H)$, where now one can think of M as being the cotangent bundle T^*Q to some configuration manifold Q. Suppose we are working with a Hamiltonian system which has a non-degenerate Poisson bracket (i.e. a symplectic manifold), then Theorem 2.15 tells us that the flow g_t^H is a symplectomorphism, and so the equations of motion are preserved, so to speak. However, there is a more powerful result that comes out of this:

Theorem 2.34 (Liouville's Theorem)**.** *Let (M, ω, H) be a Hamiltonian system of dimension $2k$. The* **volume form** *$\omega^k = \omega \wedge \cdots \wedge \omega$ is preserved under the flow. That is,*

$$\left(g_t^H\right)^* \omega^k = \omega^k. \tag{2.50}$$

Proof. Recall the following property of the pullback of a function f acting on the wedge of two forms α and β: $f^*(\alpha \wedge \beta) = f^*\alpha \wedge f^*\beta$. Then, using the result from Theorem 2.15, we have that:

$$\begin{aligned}
\left(g_t^H\right)^* \omega^k &= \left(g_t^H\right)^* (\omega \wedge \cdots \wedge \omega), \\
&= \left(g_t^H\right)^* \omega \wedge \left(g_t^H\right)^* \omega \wedge \cdots \wedge \left(g_t^H\right)^* \omega, \\
&= \omega \wedge \omega \wedge \cdots \wedge \omega, \\
&= \omega^k.
\end{aligned}$$

□

After seeing the proof, one can easily see that in fact any power of the form ω^ℓ is also preserved with our flow. The implications of Theorem 2.34 are not immediately obvious since physically it's not clear what this "volume form" does. To help with some understanding, note that ω^k can be written in local coordinates in the following way:

$$\omega^k = dq_1 \wedge \cdots \wedge dq_k \wedge dp_1 \wedge \cdots \wedge dp_k. \tag{2.51}$$

Then the integration of this form in coordinates is defined to be:

$$\int \omega^k = \int_V dq_1 \cdots dq_k dp_1 \cdots dp_k \tag{2.52}$$

where now the right hand side is the usual Riemannian integral that we have seen in calculus. Therefore, integrating ω^k gives us the integral of the volume over our whole region (we are assuming M to be compact so that this integral doesn't diverge). So what Liouville's theorem is telling us is that **phase space volume is preserved**! This has various important implications which I don't have the time to discuss, but it does have one which we will see later when dealing with the spectral stability of a Hamiltonian system – it tells us that for any exponentially growing direction, there must be an exponentially decaying one, as well. This will turn out to require that spectral stability of a Hamiltonian system is only achieved if the perturbations oscillate (they can't decay because this would imply a growing solution!). More on this in a bit.

While there are certainly a plethora of greatly important properties of Hamiltonian systems, such as other invariants, Noether's Theorem, and the issue of integrability and

chaos to name a few, I will not go into these as they do not pertain directly into what I have studied in this thesis and are not necessarily required for the understanding of my work. There are obviously plenty of references for the discussion of these things, all of which are in the many sources I have cited before: *Arnold* [1], *Audin* [2], and *Morrison* [24] are great resources.

2.2.2 Fluid Mechanics

At this point you may have noticed one fairly significant detail that I have glossed over so far – I have been dealing with finite-dimensional smooth Poisson manifolds M which describe the space on which discrete particles may live and move around. Fluid dynamics, however, is not a theory of a finite number of particles, but rather one of a *continuum* of particles, so their phase space is infinite dimensional. Now, it will turn out that most of the Hamiltonian formulation and properties that I discussed above survive in this continuum limit, but it is not something that one should consider lightly, although I will not go over the details because doing so rigorously requires far more background and is the subject of functional analysis. For the most intuitive and physically motivating transition to fluid dynamics and infinite-dimensional Hamiltonian systems, I will begin this discussion by going back to the Lagrangian description of mechanics and take the continuum limit there, after which I will motivate the equations of motion of fluid dynamics, and finally I will introduce their Hamiltonian nature. This will by *no* means be a complete depiction of fluid mechanics/dynamics, but will mostly be a description and motivation of the equations of motion and how they form a Hamiltonian system. I will not discuss any properties of fluids! It should be enough background to get a good understanding of the work that follows. For a more complete description of the subject of fluid mechanics, consult the following textbooks: for a good introduction to the subject, and one which includes a lot of intuition and experimental results to compare to theory, take a look at D. J. Tritton's *Physical Fluid Dynamics* [32]; for a more modern and serious textbook, but one that still has a good mix of rigorous theory and experimental results, consult Kundu and Cohen's *Fluid Mechanics* [14]; finally, for a more theoretical approach, but with a bit less intuition and experiments, consult Landau and Lifshitz's *Fluid Mechanics* [15]. Finally, for a thorough review of the Hamiltonian nature of the ideal fluid, see *Morrison* [24].

In the previous subsection, we began our description of particle motion by describing the position of a particle at $\vec{q}(t)$ with mass m. One can extend what I did and describe the motion of N particles by adding labels (and multiplying the dimension of the configuration space by N) to the position of each particle, say, $\vec{q}^{(i)}$ (labels will be in parenthesis whereas vector components won't be) with mass $m^{(i)}$. If we are trying to

describe a continuum of particles (i.e. a *fluid*) occupying a volume V, then we must go to a continuous label, so we will label our "fluid particles" (or "fluid parcels" or "fluid element") by their *initial position*:

$$\vec{q}(\vec{a},t) = q_i(\vec{a},t)\vec{e}_i, \tag{2.53}$$

where

$$\vec{q}(\vec{a},0) = \vec{a} \tag{2.54}$$

is their initial position and is fixed in time. What we are essentially doing is labeling eachfluid element with its initial position at $t = 0$ and then following its position $\vec{q}(\vec{a},t)$ as time evolves. Picture it as though one were following fish around an aquarium and recording its position as a function of time. Note that now each vector component $q_i(a,t)$ is no longer just a function of t, but also of the vector $\vec{a}(= a)$, so we are dealing with a configuration space Q which is really a *function space*, and hence has an infinite number of dimensions. Now anything that used to be a sum on the label i will be an integral over the label $\vec{a}$. Because we are at the continuum limit, we can no longer speak of $m^{(i)}$, but rather have to talk about the *mass density* of each particle, call it $\rho_0(\vec{a})$ – one approximation we will make is that the mass for each fluid element is conserved, so ρ_0 will *not* be a function of time. Furthermore, there are a few other complications that are added when going to the continuum limit, namely that the fluid has some sort of internal energy $U(\rho)$[9], which will be used to define the pressure eventually, given the thermodynamic definition:

$$p = \rho^2 \frac{\partial U}{\partial \rho}. \tag{2.55}$$

However, this doesn't stop us from attempting to do the same Lagrangian formalism which got us the equations of motion before. Let's begin by defining some sort of potential energy V. This potential energy will now be a *functional* rather than a function, because it must take into account all fluid elements, which requires an integral over the label $\vec{a}$ instead of a sum on i like it was before. It turns out that our potential energy is

$$V[q] = \int_V \rho_0 U(\rho_0/\mathcal{J}) d^3a, \tag{2.56}$$

where $\mathcal{J}(q)$ is the Jacobian $\mathcal{J} = \det(\partial q_i/\partial a_j)$ (I will explain this in a second). Our kinetic energy $T[q]$ is more straight forward:

$$T[q] = \frac{1}{2}\int_V \rho_0 \left(\dot{q}\right)^2 d^3a. \tag{2.57}$$

[9]Normally, U is also a function of entropy per volume, s, but we will not talk about this case. See *Morrison* [24] for a description including entropy.

And so, the Lagrangian *functional* is:

$$L[q,\dot{q}] = \int_V \left(\frac{1}{2}\rho_0 \left(\dot{q}\right)^2 - \rho_0 U(\rho_0/\mathcal{J})\right) d^3a, \tag{2.58}$$

$$= \int_V \mathcal{L}(t,q,\dot{q},\partial q/\partial a) d^3a, \tag{2.59}$$

where $\mathcal{L}$ is the **Lagrangian density**. Forgive me for the cold progression, but we are simply following the steps we have done in the finite-dimensional case. The new action functional $S[q]$ then becomes:

$$S[q] = \int_{t_0}^{t_1} \int_V \left(\frac{1}{2}\rho_0 \left(\dot{q}\right)^2 - \rho_0 U(\rho_0/\mathcal{J})\right) d^3a dt, \tag{2.60}$$

which we will now take a variation of with boundary conditions being $\delta q(a,t_0) = \delta q(a,t_1) = 0$ and that any surface terms (in terms of $\vec{a}$) vanish (this is usually a sensible thing to do, but one must be very careful with this). I will not go through the details of this variation because they require some work that will not be particularly enlightening. I recommend that the reader follow the steps in *Morrison* [24]. The resulting equation of motion for a particle is

$$\rho_0 \ddot{q}_j \frac{\partial q_j}{\partial a_i} + \mathcal{J}\frac{\partial}{\partial a_i}\left(\frac{\rho_0^2}{\mathcal{J}^2}\frac{\partial U}{\partial \rho}\right) = 0. \tag{2.61}$$

Notice the (slight) resemblance to the case with a finite number of particles. We can keep pushing the classical, finite-particle description by now defining the *canonical momentum density*:

$$\pi_i(\vec{a},t) = \frac{\delta L}{\delta \dot{q}_i(a)} = \rho_0 \dot{q}_i, \tag{2.62}$$

and getting the fluid Hamiltonian function:

$$H[q,\pi] = \int \frac{\pi^2}{2\rho_0} + \rho_0 U d^3a. \tag{2.63}$$

From this equation, we can follow the same procedure as before (but expanding on both sides in terms of δq and $\delta\pi$) and get the analogous Hamilton's equations:

$$\dot{q}_i = \frac{\delta H}{\delta \pi_i}, \tag{2.64}$$

$$\dot{\pi}_i = -\frac{\delta H}{\delta q_i}. \tag{2.65}$$

With this equation of motion, using some properties of functional derivatives (see Morrison [24] for a review), it is not hard to derive the canonical Poisson bracket bracket for a fluid:

$$\{F,G\} = \int \left(\frac{\delta F}{\delta q}\cdot\frac{\delta G}{\delta \pi} - \frac{\delta G}{\delta q}\cdot\frac{\delta F}{\delta \pi}\right) d^3a. \tag{2.66}$$

Now, we just half-derived the equation of motion with tells you how a particle at $q(a,t)$ will evolve in time, so, like in the particle description, what one is essentially doing is following the particles around as they go, seeing their position, velocity, density, and pressure change as time goes by. This is called the **Lagrangian description of a fluid**. Experimentally, though, this is not typically how one views and measures fluids. This is very challenging to do experimentally because you have to follow the 3D trajectory of a particle somehow, and there is no technique as of yet that will allow us to do this well enough. It is more practical in reality to look at one *location* in space, call it $\vec{X}$, and see how the velocity, pressure, etc. change as a function of time. This is called the **Eulerian description of a fluid**. Intuitively, this change in time of the properties at the point $\vec{X}$ will be due to not only the flow of different fluid particles with different characteristics *passing* the point $\vec{X}$ at time t, but also be due to the change in velocity or pressure that each fluid particle is experiencing as it is passing $\vec{X}$. To go from one description to the other, one must use what is called the **Lagrange-Euler map**, which maps Lagrangian variables to Eulerian ones, namely: $(\rho_0, \dot{q}) \mapsto (\rho, v)$. We begin by looking at the density, $\rho(\vec{X}, t)$:

$$\rho(\vec{X},t) = \int_V \delta(\vec{X} - \vec{q}(\vec{a},t))\rho_0(\vec{a})d^3a, \tag{2.67}$$

where $\delta(\vec{X} - \vec{q}(\vec{a},t))$ is the three-dimensional Dirac delta function. Using the following property, $\delta(f(x)) = \delta(x - x_0)/|f'(x_0)|$, where x_0 is the only place where $f(x_0) = 0$, we get a more simple expression for ρ in terms of ρ_0:

$$\rho(\vec{X},t) = \left.\frac{\rho_0(\vec{a})}{\mathcal{J}(\vec{a},t)}\right|_{\vec{a}=q^{-1}(\vec{X},t)}. \tag{2.68}$$

Eq. 2.68 is the statement of mass conservation for a fluid element – since $\mathcal{J}$ is the Jacobian determinant of the Lagrange-Euler map (and hence the volume transformation), what this is essentially saying is that the density at a point $\vec{X}$ is due to the density of the corresponding fluid particle passing that point *divided* by the change in volume of this Lagrange-Euler map. For v in terms of $\dot{q}$ we get:

$$\vec{v}(\vec{X},t) = \dot{\vec{q}}(\vec{a},t)|_{\vec{a}=q^{-1}(\vec{X},t)}. \tag{2.69}$$

Using equations 2.68 and 2.69, we can transform Eq. 2.61 into the familiar form of the equation of motion for an ideal (no viscosity, i.e. frictional forces) fluid, which, with the continuity equation, are called the **Euler equations** (with no external forces):

$$\rho\left(\frac{\partial v}{\partial t} + v\cdot\nabla v\right) = -\nabla p, \tag{2.70}$$

$$\frac{\partial \rho}{\partial t} + \nabla(\rho v) = 0. \tag{2.71}$$

Note that none of the equations of motion are closed (pressure, i.e. internal energy, aren't specified in terms of ρ or v); one usually deals with this when dealing with a more specific situation, which allows one to make some reasonable assumptions about the functional dependence of p or U. Furthermore, we don't really know where these equations live (that is, what function space), in fact this is a famous problem that is still unresolved.

Now what about the Hamiltonian structure of these Eulerian equations? Clearly the Lagrangian system is Hamiltonian, and we have simply made a coordinate change to these new equations. Therefore, the new equations must also have a Hamiltonian structure, however, it will be *non-canonical*. These equations can be described using a non-canonical Poisson bracket, which was found by Morrison and Green in 1980 [25], namely, for functionals F, G, and a fluid occupying all of $\mathbb{R}^3$ (with a velocity field that dies off at infinity allowing us to compute this integral)

$$\{F, G\} = -\int_{\mathbb{R}^3} \left(F_\rho \nabla \cdot G_{\vec{v}} + F_{\vec{v}} \cdot \nabla G_\rho - \rho^{-1}(\nabla \times \vec{v}) \cdot (F_{\vec{v}} \times G_{\vec{v}})\right) d^3X, \tag{2.72}$$

where $F_{\vec{v}} = \delta F/\delta \vec{v}$, and similarly for F_ρ. Plugging in $v, H = \int_{\mathbb{R}^3}(\rho v^2/2 + \rho U(\rho))d\vec{X}$ into this bracket gives you Eq. 2.70, and ρ, H gives you Eq. 2.71. This bracket can also be written in terms of density ρ and momentum $\vec{M} = \rho\vec{U}$:

$$\begin{aligned}\{F, G\} = \quad & -\int_{\mathbb{R}^3} M_i \left(\frac{\delta F}{\delta M_j}\frac{\partial}{\partial X^j}\frac{\delta G}{\delta M_i} - \frac{\delta G}{\delta M_j}\frac{\partial}{\partial X^j}\frac{\delta F}{\delta M_i}\right) d\vec{X} \\ & -\int_{\mathbb{R}^3} \rho \left(\frac{\delta F}{\delta M_j}\frac{\partial}{\partial X^j}\frac{\delta G}{\delta \rho} - \frac{\delta G}{\delta M_j}\frac{\partial}{\partial X^j}\frac{\delta F}{\delta \rho}\right) d\vec{X},\end{aligned} \tag{2.73}$$

where F and G are functionals of ρ and $\vec{M}$, and $\delta F/\delta M_j$ is the functional derivative of F with respect to M_j. Given the Hamiltonian $H = \int_{\mathbb{R}^3}(|\vec{M}|^2/2\rho + \rho U(\rho))d\vec{X}$, the equations

$$\dot{\rho} = \{\rho, H\}, \tag{2.74}$$

$$\dot{\vec{M}} = \{\vec{M}, H\}, \tag{2.75}$$

also generate the compressible Euler equations.

Equations 2.74 and 2.75 tell us how certain functionals (these functionals being the density and momentum components) evolve in time given their location in phase space – we call this their phase space "flow." Now, we are only interested in flows on subspaces of the phase space *where our constraints hold true*, which in our case is that the *volume is constant and the flow is divergence-free*. Note that simply setting $\nabla \cdot \vec{U} = 0$ at time $t = 0$ is not enough to ensure that the velocity field will always remain divergence-free.

Therefore, we want to restricting the phase space flow onto these subspaces of the phase space by modifying our Poisson bracket in such a way so as to make these quantities (density and divergence) constant in time (i.e. constant along our flow). This would result in the equations of motion that keep our flow on these submanifolds that we are interested in, thereby making the physics consistent with our constraints. So, leaving intuitive analogies behind in order to be precise, what we get is a new "Dirac bracket" in which the constraints (that we choose to keep constant) are *Casimirs* of the new bracket, and therefore, by equations 2.74 and 2.75, are constants of time. However, I will not do this for a general fluid first. In the next section I will go over the reduction of our specific system that we have (via the "virial method") and then Dirac constrain those reduced brackets. In the section after that I will describe the more general constrained bracket for a full 3D, incompressible fluid, but I will do so using an unconventional method developed by Chandre et al. [7].

2.2.3 Rosensteel-Dirac Bracket

For a better understanding of the physical system we are dealing with, as well as a more detailed view of the reduction of degrees of freedom that we are doing, I recommend the reader to jump straight to Chapter 3, read that, and then come back to this section for the Hamiltonian formulation of the problem. Forgive me for jumping a bit out of order, it's just that this section fits well under the Hamiltonian formulation and in theory one doesn't absolutely need Chapter 3 for what follows.

The first Hamiltonian formulation of this problem was done by Rosensteel [29] by using the "virial method" to reduce the number of degrees of freedom of the problem (look at Chapter 3 for a more complete description of this method), making the problem a system of ODE's rather than PDE's, and defining bracket relations between the virial and moment of inertia tensors. However, his formulation does not reference the fluid dynamic equations and requires the use of a non-physical energy due to the fact that the bracket is not constrained [27]. However, Morrison et al. have fixed these issues using Dirac constraint theory and so I will present their formulation here.

It turns out that applying the *full*, non-constrained, bracket in Eq. 2.73 to the virial and moment of inertia tensors, namely

$$\mathcal{M}_{ij} = \int_V \rho X_i U_j d\vec{X} = \int_V X_i M_j d\vec{X} \quad \text{and} \tag{2.76}$$

$$\Sigma_{ij} = \int_V \rho X_i X_j d\vec{X}, \tag{2.77}$$

and noting the relations

$$\frac{\delta \mathcal{M}_{ij}}{\delta M_k} = X_i \delta_{jk} \qquad \text{and} \qquad \frac{\delta \mathcal{M}_{ij}}{\delta \rho} = 0, \tag{2.78}$$

$$\frac{\delta \Sigma_{ij}}{\delta M_k} = 0 \qquad \text{and} \qquad \frac{\delta \Sigma_{ij}}{\delta \rho} = X_i X_j, \tag{2.79}$$

we get the bracket relation that was given by Rosensteel in 1988. However, these bracket relations are *not* constrained – that is, in the formulation of the original problem, we are dealing with an ellipsoid of constant volume and a flow that is incompressible, but nowhere in the Hamiltonian formulation has this been specified yet. This is precisely the reason that Rosensteel needed an unphysical Hamiltonian. We will now address that issue using Dirac constraints, defining a "Dirac bracket" which will give us the correct equations of motion generated by a physical Hamiltonian.

So we are now ready to produce the correct, constrained brackets for our system of $\mathbf{\Sigma}$ and $\mathcal{M}$. The way we will do this is not by constraining the full hydrodynamic bracket and then reducing, but rather by reducing the system first, getting the Rosensteel bracket relations and then constraining those brackets. In Section 2.2.4 I will present my original work on an attempt to approach the Rosensteel-Dirac bracket by constraining the full bracket using projections and then applying $\mathbf{\Sigma}$ and $\mathcal{M}$ to the constrained bracket, which has not been done before (spoiler: unsuccessful).

We begin by applying bracket 2.73 to equations 2.76 and 2.77, which results in the following brackets, which are the ones that Rosensteel got (which we will denote by $\{\cdot,\cdot\}_R$, for Rosensteel, like in [27]):

$$\{\Sigma_{ij}, \Sigma_{kl}\}_R = 0 \tag{2.80}$$

$$\{\mathcal{M}_{ij}, \mathcal{M}_{kl}\}_R = \delta_{il}\mathcal{M}_{kj} - \delta_{jk}\mathcal{M}_{il} \tag{2.81}$$

$$\{\mathcal{M}_{ij}, \Sigma_{kl}\}_R = -\delta_{kj}\Sigma_{li} - \delta_{lj}\Sigma_{ki}. \tag{2.82}$$

Next we want to define the Dirac bracket, but first we need to find out what our constraints are going to be. Our first constraint, which we will call C^1, is going to be the constancy of volume: $a_1 a_2 a_3 = const.$ However, we need to define C^1 in terms of our reduced variables. In order to do this, note that in the rotating frame the moment of inertia tensor is diagonal with the square of the ellipsoid axes as the diagonal elements. Therefore, since the determinant is invariant under coordinate changes, the determinant of $\mathbf{\Sigma}$ is proportional to the square of the volume of the ellipsoid. Therefore, we define, for $\tilde{\mathbf{\Sigma}}$ a dimensionless version of $\mathbf{\Sigma}$,

$$C^1 = \ln(\mathrm{Det}(\tilde{\mathbf{\Sigma}})). \tag{2.83}$$

On the other hand we have C^2, which is the constraint that says that our flow is divergence-free. For a flow that is linear in space, as is the case for us, the divergence is equal to $\mathrm{Tr}(\mathbf{K})$. But note that $\mathrm{Tr}(\mathbf{K}) = \mathrm{Tr}(\mathbf{K}^t) = \mathrm{Tr}(\mathbf{I^{-1}N}) = \mathrm{Tr}(\mathbf{\Sigma^{-1}}\mathcal{M})$. Therefore, we will define

$$C^2 = \mathrm{Tr}(\mathbf{\Sigma^{-1}}\mathcal{M}). \tag{2.84}$$

We are now ready to construct the Dirac bracket. I will post more details in the final version of this thesis, but it is pretty straight forward to get the results once you take advantage of the following relations: for a matrix $A = (A_{ij})$,

$$\frac{\partial A^{-1}_{ij}}{\partial A_{kl}} = -A^{-1}_{ik}A^{-1}_{lj}, \quad \frac{\partial \det A}{\partial A_{ij}} = C_{ij} \text{ (cofactor of } A_{ij}), \tag{2.85}$$

and, if $F_1, F_2, \ldots, F_k$ are functionals and P is a real-valued function of them, then for any other functional G,

$$\{G, P(F_1, \ldots, F_k)\} = \sum_{i=1}^{k} \frac{\partial P}{\partial F_i}\{G, F_i\}. \tag{2.86}$$

The Dirac bracket, which is the constrained version of the Rosensteel bracket and is denoted by $\{\cdot,\cdot\}_{RD}$, is as follows:

$$\{\Sigma_{ij}, \Sigma_{kl}\}_{RD} = 0 \tag{2.87}$$

$$\begin{aligned}\{\mathcal{M}_{ij}, \mathcal{M}_{kl}\}_{RD} &= \delta_{il}\mathcal{M}_{kj} - \delta_{jk}\mathcal{M}_{il} \\ &+ \frac{1}{\mathrm{Tr}\left(\mathbf{\Sigma^{-1}}\right)}\left(\delta_{ij}\left(\Sigma^{-1}_{kn}\mathcal{M}_{nl} + \Sigma^{-1}_{ln}\mathcal{M}_{nk}\right) - \delta_{kl}\left(\Sigma^{-1}_{in}\mathcal{M}_{nj} + \Sigma^{-1}_{jn}\mathcal{M}_{ni}\right)\right)\end{aligned} \tag{2.88}$$

$$\{\mathcal{M}_{ij}, \Sigma_{kl}\}_{RD} = -\delta_{kj}\Sigma_{li} - \delta_{lj}\Sigma_{ki} + \frac{2\delta_{ij}\delta_{kl}}{\mathrm{Tr}\left(\mathbf{\Sigma^{-1}}\right)}. \tag{2.89}$$

2.2.4 Alternative Derivation

In this section I will describe the general idea of using certain kinds of projectors to create Poisson brackets with Dirac constraints ("Dirac brackets"), then I will go into our specific case and describe how it will be applied to our problem.

A generic form of a Poisson bracket that we will be working with throughout this thesis[10] is (following the notation of [6]):

$$\{F, G\} = \int F_\chi \cdot \mathcal{J}(\chi) \cdot G_\chi d^n x, \tag{2.90}$$

[10] Note, this is *not* the general form of Poisson brackets in general.

where $\chi : \mathbb{R}^n \to \mathbb{R}^d$ is the field variable, $x \in \mathbb{R}^n$, and

$$F_\chi \cdot \mathcal{J}(\chi) \cdot G_\chi = F_{\chi^i} \mathcal{J}(\chi)^{ij} \cdot G_{\chi^j} = \left(\frac{\delta F}{\delta \chi^i}\right) \mathcal{J}(\chi)^{ij} \left(\frac{\delta G}{\delta \chi^j}\right).$$

We will assume that the equation of motion is given by $\dot{\chi} = \{\chi, H\}$, where $H[\chi]$ is a Hamiltonian functional. Furthermore, we will assume that $Q[\chi] = 0$ is a conservation law that is *intrinsic to the bracket of the Poisson algebra* (i.e. it doesn't depend on the Hamiltonian). This conservation law is a constraint on the system (e.g. $\nabla \cdot \vec{U} = 0$). How do we address this constraint? That is, how do we add this constraint to the Poisson algebra? There are two options for constraining: (1) We can constrain χ and then figure out how to define the functional derivatives on the constrained field variable. This turns out to be hard and is therefore a bad option. (2) We can put the constraint on the bracket, after which $Q[\chi] = 0$ takes the form of a *Casimir invariant*, which we have seen. In general physical situations it is usually the case that the constraint on the field variable χ is clearer to see and more intuitive (in our case, it would be constraining our velocity field to be divergence-free, for example), but, like I mentioned, the procedure turns out to be rather difficult. Therefore, I will go over a way to go from (1) to (2) using projectors – this procedure is summarized in [6].

So we begin with a constrained χ and an unconstrained/unaltered Poisson bracket. The idea is to "lift" this constraint on χ and move it to the Poisson bracket, but this is not a trivial thing. There are some issues – it turns out that when lifting χ to an unconstrained field, the Jacobi identity of the Poisson bracket *may no longer hold*, which is certainly a problem. Take the vorticity equation, for example. It has equations written as $\dot{\vec{\omega}} = \nabla \times (\vec{v} \times \vec{\omega})$, where $\vec{\omega} = \nabla \times \vec{v}$. Using the Poisson bracket $\{F, G\} = \int \vec{\omega} \cdot [(\nabla \times F_\omega) \times (\nabla \times G_\omega)]\, d^3x$, this becomes $\dot{\vec{\omega}} = \{\omega, H\}$. Now, in this case, the $Q[\chi] = 0$ condition is $\nabla \cdot \vec{\omega} = 0$, and our use of strategy (1) is to define $\vec{\omega} = \nabla \times \vec{v}$, so it projects/constrains the vector field to be one that is divergence-free. Now, on the quest to get to (2), if we "lift" this condition and look at the bracket we defined on a generic vector field $\vec{\omega}$, the Jacobi identity is *not* satisfied! The question then becomes, how do we take care of this? The basic strategy is to include a projector in the bracket which, once the functionals are applied, projects the fields to be divergence-free. In this specific case, we can introduce the projection operator $\mathcal{P}_\perp = (1 - \nabla \Delta^{-1} \nabla \cdot)$, which essentially takes out the non-solenoidal part of the vector field. So, for *any* vector field $\vec{u}$, $\nabla \cdot (\mathcal{P}_\perp \vec{u}) = 0$. If we now define a new Poisson bracket by $\{F, G\}' = \int (\mathcal{P}_\perp \vec{\omega}) \cdot [(\nabla \times F_\omega) \times (\nabla \times G_\omega)]\, d^3x$, we see that, even with an unconstrained field $\vec{\omega}$, this bracket satisfies the Jacobi identity. Furthermore, $\{\nabla \cdot \vec{\omega}, G\} = 0$ for all G, meaning that we have made $\nabla \cdot \vec{\omega}$ a Casimir! In this case we have reached (2). Note that using these projectors doesn't only help with

making sure that the Poisson bracket satisfies the Jacobi identity – it also is involved in the way that the functional derivatives are computed when a field variable is constrained.

Now that we have seen an example, let's go into a general method for constraining the bracket. So we are back to using bracket 2.90 with $Q[\chi] = 0$ as our constraint on the field $\chi(x)$. If χ satisfies this, then we call the field "Q-free." And again, our goal is to get the corresponding Poisson bracket for *any* field χ ($Q[\chi] = 0$ or not) which has Q as a Casimir, therefore incorporating it as a constrain on the dynamics, assuming our initial field is Q-free. We start in a situation like (1) above, where our field χ is constrained. Let $\mathcal{P}$ be a projection operator on the now *unconstrained* χ such that $\hat{Q}\mathcal{P}^\dagger = 0$, where $\hat{Q}$ is the Fréchet derivative of Q and $\mathcal{P}^\dagger$ is the adjoint of $\mathcal{P}$. It turns out (see [6] for details) that, *if* $\mathcal{P}$ doesn't depend on the field variable χ (which in our case will be true), then our constrained Poisson bracket is of the form:

$$\{F, G\}_{\mathcal{P}} = \int (\mathcal{P}F_\chi) \cdot \mathcal{J}(\mathcal{P}\chi) \cdot (\mathcal{P}G_\chi) d^n x, \tag{2.91}$$

which will satisfy the Jacobi identity for any χ and the functional $Q[\chi]$ is a Casimir. In order to choose the correct projector $\mathcal{P}$ (which is actually not unique, but there are obvious "simple" choices usually), we have to use Dirac constraint theory. We want to use Dirac constraint theory to impose a local constraint $Q[\chi](x) = 0$ (that is, make it a Casimir of the bracket). As we have seen, there is a standard procedure for doing so. The second term of the Dirac bracket can be computed to be [6]

$$-\int F_\chi \cdot \mathcal{J}\hat{Q}^\dagger (\hat{Q}\mathcal{J}\hat{Q}^\dagger)^{-1} \hat{Q}\mathcal{J} \cdot G_\chi d^n x,$$

so we can collectively call $\mathcal{J}_* := \mathcal{J} - \mathcal{J}\hat{Q}^\dagger(\hat{Q}\mathcal{J}\hat{Q}^\dagger)^{-1}\hat{Q}\mathcal{J}$ and then see that the Dirac bracket can be written as:

$$\{F, G\}_D = \int F_\chi \cdot \mathcal{J}_* \cdot G_\chi.$$

Finally, we can see that if we define $\mathcal{P}_* = 1 - \hat{Q}^\dagger(\hat{Q}\mathcal{J}\hat{Q}^\dagger)^{-1}\hat{Q}\mathcal{J}$, then we can write $\mathcal{J}_* := \mathcal{P}_*^\dagger \mathcal{J} \mathcal{P}_*$, and so, we get that

$$\{F, G\}_D = \int (\mathcal{P}_* F_\chi) \cdot \mathcal{J} \cdot (\mathcal{P}_* G_\chi). \tag{2.92}$$

This gives us a way of constraining the bracket using a projector! And, as was evident by the derivation, this projected bracket is *equivalent* to the Dirac bracket making $Q[\chi]$ a Casimir. The beauty of this method is that it simplifies the derivation of a Dirac constrained bracket. As you saw in the previous section, the Dirac constraining procedure is quite tedious. However, using this projection technique, all we need to do is find a

correct projector $\mathcal{P}$ and apply it to the unconstrained bracket that we have. Of course, we just used the Dirac constraining procedure to *find* $\mathcal{P}$, but if one somehow knows *a priori* what the projector should be (as in the example we did above with the vorticity equation), this procedure saves a lot of time and effort.

Now let's apply this theory to our specific situation: the full-on fluid dynamics bracket seen in Eq. 2.73. We first need to find out what our projector $\mathcal{P}$ is, given our constraint of choice. In our specific case, we are looking to study an incompressible fluid of uniform density. Therefore, we want our constraint to be $\nabla\cdot\vec{U} = \frac{1}{\rho}\nabla\cdot\vec{M} = 0$ (since ρ is uniform). It turns out [7] that our projector is going to be the one we saw in our previous example with the vorticity equation:

$$\mathcal{P}_\perp = 1 - \nabla\Delta^{-1}\nabla\cdot, \tag{2.93}$$

where Δ^{-1} is the three dimensional Green function of the Laplacian, defined as:

$$\Delta^{-1}f(\vec{X}) = -\frac{1}{4\pi}\int \frac{f(\vec{X}')}{|\vec{X}-\vec{X}'|}d^3X'. \tag{2.94}$$

Using Eq. 2.92 we can now write down the Dirac bracket for the fluid bracket (i.e. the incompressible fluid bracket):

$$\begin{aligned}\{F,G\}_D = \quad & \int_{\mathbb{R}^3} M_i\left(\left(\mathcal{P}_\perp\frac{\delta G}{\delta M_j}\right)\frac{\partial}{\partial X_j}\left(\mathcal{P}_\perp\frac{\delta F}{\delta M_i}\right) - \left(\mathcal{P}_\perp\frac{\delta F}{\delta M_j}\right)\frac{\partial}{\partial X_j}\left(\mathcal{P}_\perp\frac{\delta G}{\delta M_i}\right)\right)d\vec{X} \\ & + \int_{\mathbb{R}^3}\rho\left(\left(\mathcal{P}_\perp\frac{\delta F}{\delta M_j}\right)\frac{\partial}{\partial X^j}\frac{\delta G}{\delta\rho} - \left(\mathcal{P}_\perp\frac{\delta G}{\delta M_j}\right)\frac{\partial}{\partial X^j}\frac{\delta F}{\delta\rho}\right)d\vec{X}.\end{aligned} \tag{2.95}$$

What is left to do now is plug in $\mathcal{M}$ and $\mathbf{\Sigma}$ into Eq. 2.95, use the relations 2.78 and 2.79 and so on to simplify the expression and then derive the Rosensteel-Dirac brackets (equations 2.87 - 2.89). So far, after a few months of trying, this has proven to be difficult and I have not had any success. I will outline my progress below. Eq. 2.87 is easy to derive because its functional derivative with respect to $\vec{M}$ is zero, making the whole bracket zero. As for Eq. 2.89, the best I could do so far is the following:

$$\begin{aligned}\{\mathcal{M}_{ij},\Sigma_{kl}\}_{RD} \quad = \quad & -\delta_{lj}\Sigma_{ki} - \delta_{kj}\Sigma_{li} \\ & -\frac{2\delta_{jk}}{4\pi}\iint_{\mathbb{R}^3}\frac{\partial_j(\rho(\vec{X})X_i)}{|\vec{X}-\vec{X}'|}d^3\vec{X}d^3\vec{X}',\end{aligned} \tag{2.96}$$

but this does not seem to be correct, however, because I end up with an issue regarding the divergence of the integral, since we are integrating over all $\mathbb{R}^3$. There are many possible reasons for this problem, including many subtleties that we are currently looking into such as how to define the space over which we are integrating – is it over just the volume of the ellipsoid? If so, then we must change the Green's function. Or are we cutting off ρ at the boundary and integrating over all of $\mathbb{R}^3$ like Chandrasekhar did? All of this brings

up issues about non-vanishing boundary terms upon even deriving the expression in Eq. 2.95, integrating over the derivative of a discontinuous function, diverging integrals, and others. Integrating the derivative of ρ becomes tricky because we are dealing with a discontinuous function, but if we do integration by parts and then integrate over the the discontinuous ρ the integral simply becomes an integral over the volume (instead of all space like it was before). This seems to fix a few issues, but I believe that boundary terms when integrating by parts (which we have assumed always to go to zero), are not necessarily zero. This causes problems which we have not resolved yet, unfortunately. Furthermore, we need to somehow relate the Dirac bracket addition in Eq. 2.87 to that of the one we have just derived in Eq. 2.95, and there is no clear way on how to approach that at the moment. A work in progress... Now let's go back to the other major aspect of this thesis, which will be to use the correct equations of motion (seen below) for our system and look at their stability under certain specific equilibrium conditions.

2.3 Equations of Motion

With bracket relations 2.87 through 2.89 and a Hamiltonian which is in terms of $\mathbf{\Sigma}$ and $\mathcal{M}$, we can generate the equations of motion. The Hamiltonian is given in [27] as:

$$H = \frac{1}{2}\mathrm{Tr}\left(\mathcal{M}^{\mathbf{t}}\mathbf{\Sigma}^{-\mathbf{1}}\mathcal{M}\right) + \mathcal{V}(\mathbf{\Sigma}), \tag{2.97}$$

where $\mathcal{V}(\mathbf{\Sigma})$ is the gravitational potential energy (to be defined below in the rotating frame).

Note that we have been in the inertial coordinate frame in this section (hence $\vec{X}$ and not $\vec{x}$). Using the transformation $\mathbf{T}$ from Chapter 3, we can relate $\mathbf{\Sigma}$ to $\mathbf{I}$ and $\mathcal{M}$ to $\mathbf{N}$. We are interested in doing so for two reasons – due to the fact that the ellipsoid is rotating in the inertial frame, the elements of $\mathcal{M}$ and $\mathbf{\Sigma}$ will be time dependent when measured in the inertial coordinate frames, even in the equilibrium state that we will investigate later. The other reason being that everything that is formulated in the Classical Formulation section was done so in the rotating frame's coordinate system. By performing a change of coordinates in their defining integrals, we get that

$$\mathcal{M} = \mathbf{T}^{\mathbf{t}}\mathbf{N}\mathbf{T} \tag{2.98}$$

$$\mathbf{\Sigma} = \mathbf{T}^{\mathbf{t}}\mathbf{I}\mathbf{T}. \tag{2.99}$$

After plugging in these definitions to the Dirac bracket and the Hamiltonian and then doing some simple algebra, we see that, measured instantaneously in the rotating frame,

we have:

$$\{I_{ij}, I_{kl}\}_{RD} = 0 \tag{2.100}$$

$$\{N_{ij}, N_{kl}\}_{RD} = \delta_{il}N_{kj} - \delta_{jk}N_{il} + \frac{1}{\mathrm{Tr}\left(\mathbf{I^{-1}}\right)}\left(\delta_{ij}\left(I^{-1}_{kn}N_{nl} + I^{-1}_{ln}N_{nk}\right)\right. \tag{2.101}$$
$$\left. -\delta_{kl}\left(I^{-1}_{in}N_{nj} + I^{-1}_{jn}N_{ni}\right)\right)$$

$$\{N_{ij}, I_{kl}\}_{RD} = -\delta_{kj}I_{li} - \delta_{lj}I_{ki} + \frac{2\delta_{ij}\delta_{kl}}{\mathrm{Tr}\left(\mathbf{I^{-1}}\right)}, \tag{2.102}$$

and

$$H = \frac{1}{2}\mathrm{Tr}\left(\mathbf{N^t I^{-1} N}\right) + \mathcal{W}(\mathbf{I}) + \mathrm{Tr}(\mathbf{\Omega^* N}). \tag{2.103}$$

Note the change in the Hamiltonian, which is due to the fact that we are changing frames to a rotating one. In fact, this last term can also be written as the reduced angular momentum dotted with the angular velocity $\vec{\Omega}$, which is really just the reduced version of the rotational energy. Other than that, we have simply substituted one matrix for the other. The gravitational potential energy is defined to be [4] [27]

$$\mathcal{W}(\mathbf{I}) = -\frac{1}{2}\int_V \rho\mathcal{B}d\vec{x} = -\frac{8}{15}\pi^2\rho^2 a_1 a_2 a_3 G\mathcal{I}, \tag{2.104}$$

where $\mathcal{I}$ is defined in Eq. 3.6.

For the sake of time, I will not show that with bracket relations 2.100 - 2.102 and Hamiltonian 2.103 we get the right equations of motion (namely, Eq. 3.25). Look in [27] for a demonstration.

One other thing to point out is the Casimir of the original Rosensteel bracket, which will give us a total of 3 Casimirs in our Dirac bracket (two constraints and one Casimir of the original bracket). The last Casimir, let's call it C^3, is the square of the circulation [4] [27], which, in terms of $\mathbf{I}$ and $\mathbf{N}$, is:

$$C^3 := \Gamma^2 = \mathrm{Tr}\left(\mathbf{I^{-1}NIN}^t - \mathbf{NN}\right). \tag{2.105}$$

Notice that we are dealing with *matrices*, rather than the usual phase-space vector $\vec{z}$. This wouldn't necessarily be a problem, except for the fact that the equations of motion require a four-index symplectic "matrix," namely:

$$\dot{N}_{ij} = J_{ijkl}\frac{\partial H}{\partial \psi_{kl}},$$

where ψ_{kl} is a block diagonal matrix which is a combination of both $\mathbf{N}$ and $\mathbf{I}$ (so k and l would sum to up to 6). This four-index symplectic form can be cumbersome to deal with when doing calculations in *Mathematica* or even analytically. Therefore, before we

move on to the next section, I will make it so that each component of the matrices **N** and **I** are all on a single vector, $\vec{z}$, thus making the symplectic matrix a two-index object, which will make our lives easier. So, I will define a vector by placing the rows of each matrix side-by-side and then combining the vectors made from each matrix:

$$\vec{z} = (I_{11}, I_{12}, I_{13}, I_{21}, \ldots, I_{33}, N_{11}, N_{12}, \ldots, N_{33}), \tag{2.106}$$

or, equivalently,

$$\begin{aligned} I_{ij} &= z_{j+3(i-1)} \\ N_{ij} &= z_{j+3(i-1)+9}. \end{aligned} \tag{2.107}$$

So now our equation of motion is of the form

$$\dot{z}_\alpha = J_{\alpha\beta}\frac{\partial H}{\partial z_\beta}, \tag{2.108}$$

which is what we wanted. The convention from now on will be that greek letters are summed from 1 to 18 whereas latin letters are summed from 1 to 3. Now the question becomes - what is $J_{\alpha\beta}$? Note that, with the Dirac bracket we have defined, the equations of motion are $\dot{z}_\alpha = \{z_\alpha, H\}_{RD}$. However, if we use Eq. 2.86, we get:

$$\dot{z}_\alpha = \{z_\alpha, H\}_{RD} = \{z_\alpha, z_\beta\}_{RD}\frac{\partial H}{\partial z_\beta}, \tag{2.109}$$

noting again that we sum β from 1 to 18. Therefore, we can now define $J_{\alpha\beta}$:

$$J_{\alpha\beta} = \{z_\alpha, z_\beta\}_{RD}. \tag{2.110}$$

We are now ready to look at the equilibria and stability of the Riemann Ellipsoids using our Hamiltonian formulation. But first we will study what I call the "classical formulation" of this problem, which includes the virial method done by Chandrasekhar and Lebovitz.

Chapter 3

Classical Formulation

Following Chapter 4 of Chandrasekhar's book [4], we will denote coordinates of the inertial frame by X_i and those in the rotating frame by x_i, for $i = 1, 2, 3$. Let $\mathbf{T}(t)$ be the coordinate transformation that transforms the inertial frame coordinate system into that of the rotating frame:

$$\vec{x} = \mathbf{T}(t)\vec{X} \qquad \text{or} \qquad x_i = T_{ij}(t)X_j. \tag{3.1}$$

(Henceforth I will not explicitly denote the time dependence of $\mathbf{T}$.) Now, it follows from the fact that $\mathbf{T}$ is orthogonal that, the matrix

$$\mathbf{\Omega}^* = \frac{d\mathbf{T}}{dt}\mathbf{T}^t$$

is anti-symmetric and therefore is of dimension 3. With this matrix we can define the rotation vector $\vec{\Omega}$ of the rotating frame with respect to the inertial frame, by

$$\Omega^*_{ij} = \epsilon_{ijk}\Omega_k.$$

I will denote the inertial frame velocity *measured in the rotating frame's coordinate system* by $\mathbf{T}d\vec{X}/dt = \vec{U}$ and the rotating frame velocity $d\vec{x}/dt = \vec{u}$. It can be shown that

$$\vec{U} = \vec{u} - \mathbf{\Omega}^*\vec{x} = \vec{u} + \vec{\Omega} \times \vec{x}, \tag{3.2}$$

$$\mathbf{T}\frac{d^2\vec{X}}{dt^2} = \frac{d\vec{U}}{dt} - \mathbf{\Omega}^*\vec{U}, \tag{3.3}$$

where on the left hand side of Eq. 3.2 is the *inertial velocity measured instantaneously in the rotating coordinates* and in the last equation, Eq. 3.3, is the *acceleration in the inertial frame, measured instantaneously in the rotating coordinates.*

We are studying the motions of an ellipsoid with constant volume, divergence-free flow, and with time-dependent semi-axes, a_1, a_2, and a_3. Without loss of generality, we will choose a_i to align with x_i, the coordinate axes in the frame of reference of the rotating ellipsoid (which is rotating with angular velocity $\vec{\Omega}$). The acceleration in the inertial frame, but measured in the rotating frame, is:

$$\rho\frac{d\vec{U}}{dt} - \rho\mathbf{\Omega}^*\vec{U} = -\nabla p + \rho\nabla\mathcal{B}, \tag{3.4}$$

where $\mathcal{B}$ is the gravitational potential inside a homogeneous ellipsoid. Note that the gradient is evaluated in the coordinates of the moving frame and that

$$\frac{d\vec{U}}{dt} = \frac{\partial\vec{U}}{\partial t} + \vec{u}\cdot\nabla\vec{U}.$$

Now, $\mathcal{B}$ is given by[4]:

$$\mathcal{B} = \pi G\rho\left(\mathcal{I} - \sum_{i=1}^{3}\mathcal{A}_i x_i^2\right), \tag{3.5}$$

where,

$$\mathcal{I} = a_1a_2a_3\int_0^\infty \frac{du}{\Delta(u)}, \tag{3.6}$$

$$\mathcal{A}_i = a_1a_2a_3\int_0^\infty \frac{du}{\Delta(u)(a_i^2+u)}, \tag{3.7}$$

and

$$\Delta(u) = \sqrt{(a_1^2+u)(a_2^2+u)(a_3^2+u)}. \tag{3.8}$$

We may now begin to describe the moment reduction, otherwise known in this case as the *second order virial equations*. We begin by multiplying the jth component of Eq. 3.4 by x_i (to be consistent with [27]) and integrating over the volume of the ellipsoid $V = (4/3)\pi a_1a_2a_3$.

$$\frac{d}{dt}\int_V \rho x_i U_j d\vec{x} = \int_V \rho u_i U_j d\vec{x} + \int_V \rho\Omega^*_{jn}x_iU_n d\vec{x} + \Pi\delta_{ij} + \mathfrak{B}_{ij}, \tag{3.9}$$

where

$$\Pi\delta_{ij} = -\int_V x_j\frac{\partial p}{\partial x_i}d\vec{x} = \delta_{ij}\int_V p\; d\vec{x}, \tag{3.10}$$

and

$$\mathfrak{B}_{ij} = \int_V \rho x_i\frac{\partial\mathcal{B}}{\partial x_j}d\vec{x} = -G\int_V\int_{V'}\rho(\vec{x})\rho(\vec{x}')x_i'\frac{x_j - x_j'}{|\vec{x}-\vec{x}'|^3}d\vec{x}d\vec{x}'. \tag{3.11}$$

Let's look at the case of when $\vec{U}$ is a linear function of $\vec{x}$, and so we can define the relation as:

$$\vec{U} = \mathbf{K}\vec{x} \quad \text{or} \quad U_i = K_{ij}x_j. \tag{3.12}$$

Plugging in Eq. 3.12 in Eq. 3.9, and defining the moment of inertia tensor

$$I_{ij} = \int_V \rho x_i x_j d\vec{x}, \tag{3.13}$$

we get:

$$\frac{d}{dt}(K_{jl}I_{il}) = K_{il}K_{jn}I_{ln} + K_{jl}\Omega^*_{in}I_{nl} + K_{nl}\Omega^*_{jn}I_{il} + \Pi\delta_{ij} + \mathfrak{B}_{ij}, \tag{3.14}$$

which, using the fact that Ω^* is anti-symmetric, we can write as,

$$\frac{d}{dt}(\mathbf{I}\mathbf{K}^t) = \mathbf{K}\mathbf{I}\mathbf{K}^t + \mathbf{\Omega}^*\mathbf{I}\mathbf{K}^t - \mathbf{I}\mathbf{K}^\mathbf{t}\mathbf{\Omega}^* + \Pi\,\mathbb{I} + \mathfrak{B}. \tag{3.15}$$

Furthermore, looking at Eq. 3.5, we see that $(\nabla\mathcal{B})_i = -2\pi G\rho\mathcal{A}_i x_i$ (no sum). And so, this means that our final equation is:

$$\frac{d}{dt}(\mathbf{I}\mathbf{K}^t) = \mathbf{K}\mathbf{I}\mathbf{K}^t + \mathbf{\Omega}^*\mathbf{I}\mathbf{K}^t - \mathbf{I}\mathbf{K}^\mathbf{t}\mathbf{\Omega}^* + \Pi\,\mathbb{I} - \frac{3}{2}mG\mathcal{A}\mathbf{I}, \tag{3.16}$$

where now $\mathcal{A} = \text{diag}(\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3)$. It is now convenient to define the matrix $\mathbf{N} = \mathbf{I}\mathbf{K}^t$, which will be used later. This is the "virial tensor" in the rotating coordinate frame. Eq. 3.16 is now:

$$\frac{d}{dt}\mathbf{N} = \mathbf{N}^\mathbf{t}\mathbf{I}^{-\mathbf{1}}\mathbf{N} + \mathbf{\Omega}^*\mathbf{N} - \mathbf{N}\mathbf{\Omega}^* + \Pi\,\mathbb{I} - 2\pi\rho G\mathcal{A}\mathbf{I}. \tag{3.17}$$

In fact, using Eq. 2.104, we can write Π as [27]:

$$\Pi = \frac{1}{\text{Tr}(\mathbf{I}^{-1})}\text{Tr}\left((\mathbf{N}^\mathbf{t}\mathbf{I}^{-\mathbf{1}})^2 + 2\frac{\partial\mathcal{W}}{\partial\mathbf{I}}\right). \tag{3.18}$$

We would like to find out a bit more about $\mathbf{K}$, since this is the matrix that determines the time-dependence of $\vec{U}$. We will do so by looking at the so called "Riemann-Lebovitz formulation of the Dirichlet Problem." We begin by assuming two things: the constancy of our volume, or equivalently $a_1a_2a_3 = const$, and the fact that the position of a fluid element in the inertial reference frame, $\vec{X}(t)$, *is a linear function of* $\vec{x}(0)$:

$$\vec{X}(t) = \mathbf{P}(t)\mathbf{A}^{-1}(0)\vec{x}(0) \quad \text{or} \quad X_i(t) = P_{ij}(t)\frac{x_j(0)}{a_j(0)}, \tag{3.19}$$

where

$$A_{ij} = \delta_{ij}a_j \qquad \text{(no sum)}. \tag{3.20}$$

If we now multiply on both sides by $\mathbf{A^{-1}T}$, we get:

$$\mathbf{A}^{-1}\vec{x}(t) = \mathbf{A^{-1}TPA^{-1}}(0)\vec{x}(0) := \mathbf{S}(t)\mathbf{A}^{-1}(0)\vec{x}(0).$$

This tells us that $\mathbf{S}$ essentially gives us the internal dynamics - that is, how $\vec{x}(t)$ evolves, but also taking into account the evolution of the semi-axes a_i's. It turns out (see [4], page 68) that $\mathbf{S}$ is in fact an orthogonal transformation, which means we can define another "rotation" vector (which will turn out to be proportional to the vorticity of the internal motion):

$$\mathbf{\Lambda}^* = \frac{d\mathbf{S}}{dt}\mathbf{S}^t, \tag{3.21}$$

with the vector defined, similarly to $\vec{\Omega}$, as $\Lambda^*_{ij} = \epsilon_{ijk}\Lambda_k$.

And so, Dirichlet's problem is determining whether a matrix $\mathbf{P}$ exists and, at the same time, is expressible in terms of two orthogonal matrices $\mathbf{T}$ and $\mathbf{S}$ by $\mathbf{P} = \mathbf{T^tAS}$. We can determine this by using the equations of motion of an incompressible fluid and expressing them in terms of these matrices. We begin by looking at $\vec{U}$, which, when measured in the rotating frame, is defined as $\vec{U} = \mathbf{T}\frac{d\vec{X}}{dt}$, so we get:

$$\begin{aligned}
\vec{U} &= \mathbf{T}\frac{d\vec{X}}{dt} \\
&= \mathbf{T}\frac{d\mathbf{P}}{dt}\mathbf{A}^{-1}(0)\vec{x}(0) \\
&= \mathbf{T}\frac{d\mathbf{P}}{dt}\mathbf{S}^t\vec{x}(t).
\end{aligned}$$

But note that $\mathbf{P} = \mathbf{T^tAS}$, so that means that,

$$\begin{aligned}
\vec{U}(t) &= \mathbf{T}\frac{d\mathbf{T^tAS}}{dt}\mathbf{S}^t\vec{x}(t) \\
&= \mathbf{T}(\mathbf{T}^t\mathbf{A}\frac{d\mathbf{S}}{dt} + \frac{\mathbf{T}^t}{dt}\mathbf{AS} + \mathbf{T}^t\frac{\mathbf{A}}{dt}\mathbf{S})\mathbf{S}^t\mathbf{A}^{-1}\vec{x}(t) \\
&= \left(\mathbf{A\Lambda}^*\mathbf{A}^{-1} - \mathbf{\Omega}^* + \frac{d\mathbf{A}}{dt}\mathbf{A}^{-1}\right)\vec{x}(t).
\end{aligned}$$

So,

$$\vec{U}(t) = \left(\mathbf{A\Lambda}^*\mathbf{A}^{-1} - \mathbf{\Omega}^* + \frac{d\mathbf{A}}{dt}\mathbf{A}^{-1}\right)\vec{x}(t). \tag{3.22}$$

And there we have it:

$$\mathbf{K} = \mathbf{A\Lambda}^*\mathbf{A}^{-1} - \mathbf{\Omega}^* + \frac{d\mathbf{A}}{dt}\mathbf{A}^{-1}. \tag{3.23}$$

One result of this is that we get a more physical meaning for Λ^*. By looking at each component of $\vec{U}$ we realize that

$$\omega_k = -\frac{a_i^2 + a_j^2}{a_i a_j}\Lambda_k, \qquad (i \neq j \neq k),$$

where $\vec{\omega}$ is the vorticity vector of the fluid flow *in the rotating frame* defined by $\nabla \times \vec{u}$. Another thing to notice is that, for an ellipsoid, the moment of inertia tensor is diagonal and given by

$$\mathbf{I} = \frac{m}{5}\mathbf{A}^2. \tag{3.24}$$

Plugging the results of equations 3.23 and 3.24 back into Eq. 3.16 and simplifying we get:

$$\frac{d^2\mathbf{A}}{dt^2} + \frac{d}{dt}\left(\mathbf{A}\boldsymbol{\Lambda}^* - \boldsymbol{\Omega}^*\mathbf{A}\right) + \frac{d\mathbf{A}}{dt}\boldsymbol{\Lambda}^* - \boldsymbol{\Omega}^*\frac{d\mathbf{A}}{dt} + \mathbf{A}\boldsymbol{\Lambda}^{*2} + \boldsymbol{\Omega}^{*2}\mathbf{A} - 2\boldsymbol{\Omega}^*\mathbf{A}\boldsymbol{\Lambda}^* = \frac{2p_c}{\rho}\mathbf{A}^{-1} - 2\pi\rho G\mathcal{A}\mathbf{A}. \tag{3.25}$$

In order to be consistent with [27], it is worth mentioning that $\frac{2p_c}{\rho} = \frac{5\ \Pi}{m}$, where p_c is the pressure at the center of the ellipsoid.

Equation 3.25 gives us the time dependence of the three axes, the three components of the rotation vector, the three components of the vorticity-like vector, and the pressure – making a total number of 10 unknowns.

Chapter 4

Equilibria and Stability Analysis

4.1 Equilibrium Conditions

We are interested in finding the equilibrium configurations of our system. This means that we are looking for points (called z^e) such that

$$\dot{z}_\alpha = \{z_\alpha, H\} = 0.$$

In canonical Hamiltonian systems, one can write this expression as

$$\dot{z}_\alpha = J_c^{ij} \frac{\partial H}{\partial z_\alpha} = 0,$$

and since $\det J_c \neq 0$, the requirement is simply that

$$\left.\frac{\partial H}{\partial z_\alpha}\right|_{z^e} = 0.$$

However, in non-canonical systems (like our own), $\det J = 0$, so while $\partial H/\partial z_\alpha = 0$ is *a* solution, it is not the only possible equilibrium solution (in fact, it tends to give trivial/uninteresting solutions such as $\vec{v} = 0$). This is because we are not considering the constraints on our system given by our Casimir invariants – that is, we are not considering the fact that we are *constrained* to lie on a symplectic leaf in our Poisson manifold, and so our expression for finding an extremum of our system is not in general correct. In order to find the correct extrema of our system, we must consider the Casimir invariants and use *Lagrange multipliers*, so we will instead be minimizing what we call the **Energy-Casimir Function**, denoted by F, and defined to be:

$$F = H + c_i C^i, \tag{4.1}$$

where H is the Hamiltonian of Eq. 2.103, C^i's are the Casimir functions, and c_i's are Lagrange multipliers.

What follows is also my original work; re-deriving the equilibrium family of ellipsoids which we are studying using our Hamiltonian formulation, and then an attempt to study the stability of these ellipsoids via spectral stability methods as well as nonlinear ones such as Theorem 1.3.

Our system has the following Energy-Casimir:

$$F = \frac{1}{2}\mathrm{Tr}\left(\mathbf{N^t I^{-1} N}\right)+\mathcal{W}(\mathbf{I})+\mathrm{Tr}(\mathbf{\Omega^* N})+c_1\ln(\mathrm{Det}(\tilde{\mathbf{I}}))+c_2\mathrm{Tr}(\mathbf{I^{-1}N})+c_3\mathrm{Tr}\left(\mathbf{I^{-1}NIN}^t-\mathbf{NN}\right). \tag{4.2}$$

After taking the gradient of F, I got the following:

$$\begin{aligned}\frac{\partial F}{\partial I_{ij}} = \frac{\partial F}{\partial z_{j+3(i-1)}} &= -\frac{1}{2}I^{-1}_{ki}N_{ks}N_{ls}I^{-1}_{jl}+\frac{\partial \mathcal{W}}{\partial I_{ij}}+c_1 I^{-1}_{ij}-c_2 N_{rs}I^{-1}_{si}I^{-1}_{jr}+ \\ &\quad c_3\left(I^{-1}_{ts}N_{si}N_{tj}-N_{sr}I_{rn}N_{tn}I^{-1}_{ti}I^{-1}_{js}\right), \end{aligned} \tag{4.3}$$

$$\begin{aligned}\frac{\partial F}{\partial N_{ij}} = \frac{\partial F}{\partial z_{j+3(i-1)+9}} &= \frac{1}{2}\left(I^{-1}_{ik}N_{kj}+I^{-1}_{ri}N_{rj}\right)-\Omega^*_{ij}+c_2 I^{-1}_{ij}+ \\ &\quad c_3\left(I^{-1}_{is}N_{sr}I_{rj}+I^{-1}_{ti}N_{tn}I_{jn}-2N_{ji}\right), \end{aligned} \tag{4.4}$$

where

$$\frac{\partial \mathcal{W}}{\partial I_{ij}} = \begin{cases} \rho\pi G\mathcal{A}_i & \text{for } i=j, \\ 0 & \text{for } i\neq j. \end{cases} \tag{4.5}$$

Eq. 4.5 can be derived by noting that

$$\frac{\partial \mathcal{W}}{\partial I_{ij}} = \frac{5}{m}\frac{\partial \mathcal{W}}{\partial a_i^2}$$

and then taking the derivative of $\mathcal{W}$ with respect to a_i^2, noting that $a_1a_2a_3$ (proportional to the volume) is constant.

Before we go on, it is important to point out once more that we will be looking at the specific set of equilibrium conditions where *the axes of rotation and vorticity coincide with a principle axis of the ellipsoid,* which, without loss of generality, we will choose to be a_3 (this is the case of equilibria studied in [3]). Like I mentioned in Chapter 1, it turns out that the set of equilibrium ellipsoids is a two-parameter family, given by Eq. (38) in [3], with the parameters being $f := \omega_3/\Omega_3$ and $\tilde{a}_3 := a_3/a_1$. Furthermore, we were interested in studying one specific case of f, namely:

$$f = -\frac{a_1^2+a_2^2}{a_1a_2}. \tag{4.6}$$

This then gives us a one-parameter family of equilibria, which make the analysis a bit more simple. These conditions, and noting that $f = -a_1^2 + a_2^2/a_1 a_2$ implies that $\mathbf{\Omega}^* = \mathbf{\Lambda}^*$, most noticeably change the matrix $\mathbf{K}$ (see Eq. 3.23), which, at equilibrium, becomes:

$$\mathbf{K} = \begin{pmatrix} 0 & \frac{a_1\Omega_3}{a_2} - \Omega_3 & 0 \\ \Omega_3 - \frac{a_2\Omega_3}{a_1} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.7}$$

After calculating the gradient of F, applying our equilibrium assumptions (that is, aligning $\vec{\Omega}, \vec{\omega}$, and a_3, and also taking $f = -(a_1^2 + a_2^2)/(a_1 a_2)$), and setting it all equal to the zero vector, I was able to solve for the Lagrange multipliers and got:

$$c_1 = -\frac{4}{15} a_1 a_2 a_3^3 (\pi\rho)^2 G \mathcal{A}_3, \tag{4.8}$$

$$c_2 = 0, \tag{4.9}$$

$$c_3 = -\frac{15}{8\pi\rho(a_1 - a_2)^2 a_1 a_2 a_3}. \tag{4.10}$$

Furthermore, I was able to solve for the equilibrium value of Ω_3:

$$\Omega_3 = \sqrt{\rho\pi G \left(\frac{-a_2^2\mathcal{A}_2 + a_3^2\mathcal{A}_3}{a_1 a_2 - a_2^2} \right)}.$$

We can simplify this expression (and you'll see why we want to in a second) using equations (104) and (106) of Chapter 3 of [4] to show that

$$\frac{-a_2^2\mathcal{A}_2 + a_3^2\mathcal{A}_3}{a_1 a_2 - a_2^2} = \frac{a_3^2 - a_2^2}{a_2(a_1 - a_2)} B_{23},$$

where

$$B_{ij} = \int_0^\infty \frac{u du}{\Delta(u)(a_i^2 + u)(a_j^2 + u)}. \tag{4.11}$$

Next, using equations (24) and (26) of [3] we can simplify further and show that

$$\frac{a_3^2 - a_2^2}{a_2(a_1 - a_2)} B_{23} = B_{12},$$

which gives us the final expression for Ω_3:

$$\Omega_3 = \sqrt{\rho\pi G B_{12}}. \tag{4.12}$$

With these Lagrange multipliers, I was left with one final expression relating $\tilde{a}_3$ and $\tilde{a}_2 := a_2/a_1$, giving us the shape of the equilibrium ellipsoids in our family. The resulting the relationship can be seen in Figure 4.1. This plot, along with the assumption that $a_1 > a_2$ which is made in [3], tells us that this family of equilibrium ellipsoids are all

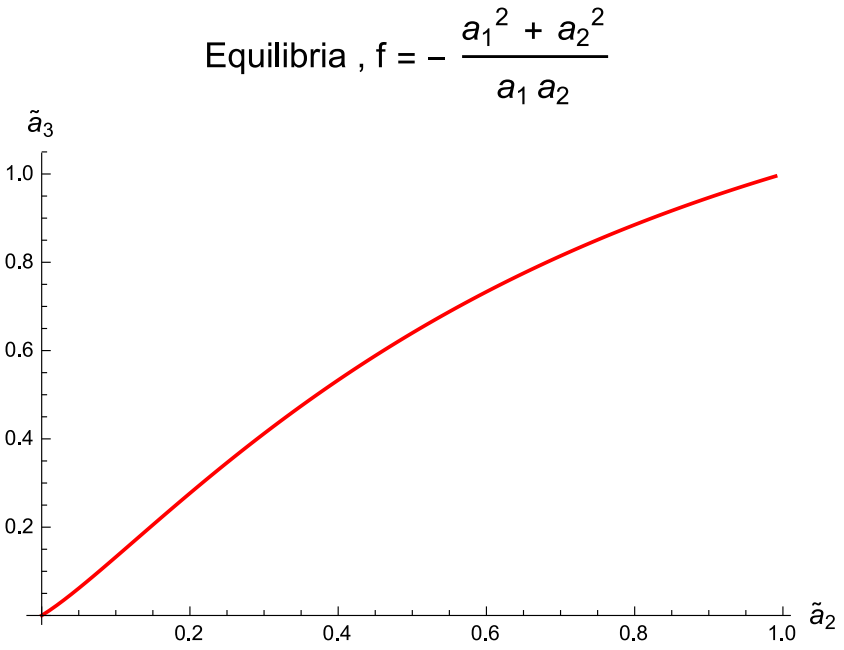


FIGURE 4.1: The values of $\tilde{a}_2 := a_2/a_1$ as a function of the parameter $\tilde{a}_3$ that give equilibrium ellipsoids for our given value of f. Note that $\tilde{a}_2$ is given in the x-axis. This is done so that one can compare to Figure 2 in [3]. This plot, along with the assumption that $a_1 > a_2$ which is made in [3], tells us that this family of equilibrium ellipsoids are all *prolate*, meaning that $a_1 > a_3 > a_2$.

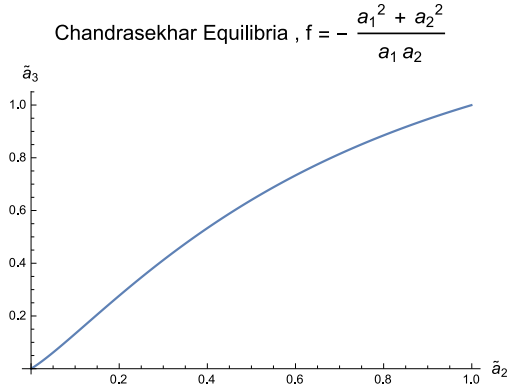
prolate, meaning that $a_1 > a_3 > a_2$ (see Figure 1.1).

We are back, for a moment, to the notation of the classical formulation with which Chandrasekhar did his work. Our goal is to confirm that our equilibria ellipsoids match that of Chandrasekhar's. As we saw in Chapter 3, the dynamics of our ellipsoid axes, the solid body rotation, and the internal flow are all determined explicitly by Eq. 3.25. Therefore, we will be looking at the equilibrium conditions given by that equation. Figure 4.2 shows the resulting roots of Eq. (38) from Chandrasekhar's paper and the comparison of the two. Furthermore, Chandrasekhar tells us that these equilibrium values of $\tilde{a}_2$ and $\tilde{a}_3$ correspond to the following value of Ω_3:

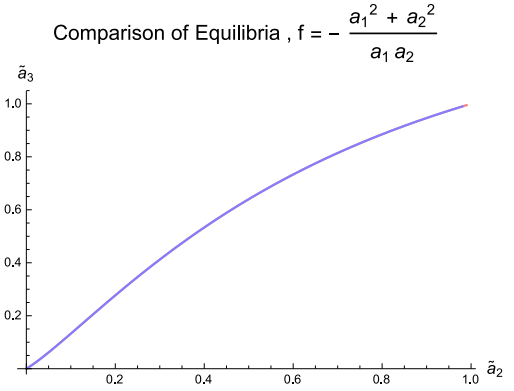
$$\Omega_3^2 = \rho\pi G B_{12}, \tag{4.13}$$

which is exactly what we got.

For the computational details, including the Mathematica notebook as it is, please refer to Section A.1 of Appendix A.



(a) The values of $\tilde{a}_2 := a_2/a_1$ as a function of the parameter $\tilde{a}_3$ that give equilibrium ellipsoids for our given value of f. Calculated from Eq. (38) from [3].



(b) A comparison of Chandrasekhar's equilibrium results given by Eq. 3.25 (in red) and our results given by the extermization of F, i.e. Eq. 4.2 (in blue).

FIGURE 4.2: As you can see, these values agree, thereby giving further confirmation that our formulations are equivalent.

4.2 Stability

Now that we have established our equilibrium conditions, we are ready to move on to stability calculations of said equilibria. First I will go over the spectral stability of our system and compare the results to that of Chandrasekhar's, which should agree. Then, I will comment on the nonlinear stability of our system with respect to Theorem 1.3.

The idea behind spectral stability is to linearize the equation of motion around an equilibrium point, assume the perturbation δz is of the form $\delta z(t) = \delta\hat{z}e^{\lambda t}$, plug this form into the linearized equations of motion, and then study the resulting eigenvalue problem, which, computationally speaking, is very doable and easy to do. That is, we will substitute $z^e + \delta z$ for z in Eq. 2.108, for small $\delta z(t)$ and z^e the equilibrium value (which does not depend on time), and then simplify using the equilibrium conditions and by ignoring every term that has a power of δz which is greater than one. This results in the linearized equation, which can be easily shown to be

$$\delta\dot{z}_\alpha = J^e_{\alpha\nu}A^e_{\nu\beta}\delta z_\beta =: \mathcal{H}^e_{\alpha\beta}\delta z_\beta, \tag{4.14}$$

where $A^e_{\nu\beta} = \partial^2 F/\partial z_\nu \partial z_\beta|_e$ is the Hessian matrix of F evaluated at an equilibrium value, and the e represents the fact that we are evaluating this matrix at an equilibrium point $(\tilde{a}_2(\tilde{a}_3), \tilde{a}_3)$ corresponding to $f = -(a_1^2 + a_2^2)/(a_1 a_2)$. Now, the spectral stability of this linearized system is determined by assuming $\delta\dot{z} = \lambda\delta\hat{z}e^{\lambda t}$ and looking at the eigenvalues of $\mathcal{H}^e$ – stability referring to the situation in which *all* of the eigenvalues have a real part equal to zero (but they can be exactly equal to zero, which is the case for many eigenvalues, in fact, as Chandrasekhar showed).

As for the question of nonlinear stability (following the argument given by Salmon in [30]), consider small perturbations δz from the equilibrium point z^e, which is an extremum of the Energy-Casimir F. We define

$$\begin{aligned} \Delta F(\delta z; z^e) &= F(z^e + \delta z) - F(z^e), \\ &= \frac{1}{2}\frac{\partial^2 F}{\partial z^\alpha \partial z^\beta}\delta z^\alpha \delta z^\beta + O(\delta z^3). \end{aligned} \tag{4.15}$$

A nonlinearly *stable* equilibrium state is one where the quadratic form in Eq. 4.15 is definite. This is in fact the same statement as Theorem 1.3, which states:

> Let the Hessian matrix $\frac{\partial^2 F}{\partial z^\alpha \partial z^\beta}$ of the Energy-Casimir F be definite at an equilibrium z^e. Then z^e is stable.

Therefore, all that we must do is look at the signs of the eigenvalues of the Hessian matrix of F. The intuition behind this is the following: imagine we are dealing with $F : \mathbb{R}^2 \to \mathbb{R}$ and picture its graph in $\mathbb{R}^3$; if the quadratic form (hessian matrix) is definite, then that means that all of its eigenvalues are of the same sign. The eigenvalues of the the hessian matrix tell us the *principal curvatures*, at the equilibrium point, of the surface that is the graph of F (on which our dynamics lie, following the analogy). If they are all of the same sign this means that we are dealing with an elliptic paraboloid, and *not* saddle-like surface, giving us stability for finite-size perturbations δz.

It is important to once again mention what I said in the intro: we are dealing with a *reduced* system, and even though it describes the dynamics exactly, the stability of this reduced system is not exactly the same as the stability of the real, infinite-dimensional system! However, if we find that our reduced system is unstable, then the full system must also be unstable; but if we find that our reduced system is stable, this does not necessarily mean that our full system is stable. We will not be tackling this question in this thesis, but some stability calculations have been done for our type of ellipsoids – see Lebovitz and Lifschitz's papers [19, 20].

Finally, even though we won't be directly applying this method in this thesis, it is worth saying a word or two about another alternative to study the stability of our system, namely the *Rayleigh-Ritz* variational method. In general infinite-dimensional Hamiltonian systems, one can use a very similar analog to Theorem 1.3, where one has to prove that the second variation of the Energy-Casimir (now a functional) is positive with respect to any arbitrary variation [24]. This is generally very hard to prove, but for Hamiltonian systems which undergo a steady state bifurcation (see below) one can use what is called the *Rayleigh-Ritz* variational method [13]. In this method, one only needs to find *trial functions* which make the Rayleigh quotient negative, thereby showing

instability. The only issue is that this would not give us the exact parameter value at which the instability occurs, but it would at least be nice to be able to say something about the nonlinear stability of our system in certain parameter regions. In particular, showing instability at any parameter value in which the system is spectrally stable would be a very important find.

And so, we are ready to look at the stability of our system. What is left for us to do, since we have already calculated the matrix $J_{\alpha\nu}$ (Eq. 2.110), is to calculate the Hessian matrix $A^e_{\nu\beta}$. We will do so by taking the derivative of the gradient, given by Eq. 4.3. Note that, even though the Lagrangian multipliers are functions of a_i, we will still treat them as constants since these Lagrange multipliers are simply supposed to be evaluated at the equilibrium point of interest, and are not part of the general F. Here are the results that I have calculated analytically:

$$\begin{aligned}\frac{\partial}{\partial I_{ab}}\left(\frac{\partial F}{\partial I_{ij}}\right) &= \frac{\partial^2 F}{\partial z_{b+3(a-1)}\partial z_{j+3(i-1)}} = \frac{1}{2}N_{ks}N_{ls}I^{-1}_{jl}I^{-1}_{ka}I^{-1}_{bi} + \frac{1}{2}N_{ks}N_{ls}I^{-1}_{ki}I^{-1}_{ja}I^{-1}_{bl} \\ &+\frac{\partial^2 \mathcal{W}}{\partial I_{ab}\partial I_{ij}} - c_1 I^{-1}_{ia}I^{-1}_{bj} + c_3\Big(N_{sr}N_{tn}I_{rn}I^{-1}_{ti}I^{-1}_{ja}I^{-1}_{bs} + N_{sr}N_{tn}I_{rn}I^{-1}_{js}I^{-1}_{ta}I^{-1}_{bi} \\ &-N_{sa}N_{tb}I^{-1}_{ti}I^{-1}_{js} - N_{si}N_{tj}I^{-1}_{ta}I^{-1}_{bs}\Big)\end{aligned} \tag{4.16}$$

$$\begin{aligned}\frac{\partial}{\partial I_{ab}}\left(\frac{\partial F}{\partial N_{ij}}\right) &= \frac{\partial^2 F}{\partial z_{b+3(a-1)}\partial z_{j+3(i-1)+9}} = -\frac{1}{2}N_{kj}I^{-1}_{ia}I^{-1}_{bk} - \frac{1}{2}N_{rj}I^{-1}_{ra}I_{bi} \\ &+c_3\left(N_{sa}I^{-1}_{is}\delta_{jb} - N_{sr}I_{rj}I^{-1}_{ia}I^{-1}_{bs} + N_{tb}I^{-1}_{ti}\delta_{ja} - N_{tn}I_{jn}I^{-1}_{ta}I^{-1}_{bi}\right)\end{aligned} \tag{4.17}$$

$$\begin{aligned}\frac{\partial}{\partial N_{ab}}\left(\frac{\partial F}{\partial I_{ij}}\right) &= \frac{\partial^2 F}{\partial z_{b+3(a-1)+9}\partial z_{j+3(i-1)}} = -\frac{1}{2}N_{kb}I^{-1}_{ki}I^{-1}_{ja} - \frac{1}{2}N_{lb}I^{-1}_{ai}I^{-1}_{jl} \\ &+c_3\Big(-I_{rb}I^{-1}_{ai}I^{-1}_{js}N_{sr} - I_{bn}I^{-1}_{ti}I^{-1}_{ja}N_{tn} + I^{-1}_{as}N_{si}\delta_{jb} + I^{-1}_{ta}N_{tj}\delta_{ib}\Big)\end{aligned} \tag{4.18}$$

$$\begin{aligned}\frac{\partial}{\partial N_{ab}}\left(\frac{\partial F}{\partial N_{ij}}\right) &= \frac{\partial^2 F}{\partial z_{b+3(a-1)+9}\partial z_{j+3(i-1)+9}} = \frac{1}{2}I^{-1}_{ia}\delta_{jb} + \frac{1}{2}I^{-1}_{ai}\delta_{jb} + \\ &c_3\left(I^{-1}_{ia}I_{bj} + I^{-1}_{ai}I_{jb} - 2\delta_{aj}\delta_{bi}\right),\end{aligned} \tag{4.19}$$

where

$$\frac{\partial^2 \mathcal{W}}{\partial I_{ab}\partial I_{ij}} = \begin{cases} 0 & \text{if } a \neq b \text{ or } i \neq j, \\ -\frac{45}{8}G\int_0^\infty \frac{du}{\Delta(a_i^2+u)^2} & \text{if } a = b = i = j, \\ -\frac{15}{8}G\int_0^\infty \frac{du}{\Delta(a_i^2+u)(a_j^2+u)} & \text{if } a = b \neq i = j. \end{cases} \tag{4.20}$$

The transition (as we change the parameter) from stable to unstable, and vice versa, is called a *bifurcation*[1]. Before I go into the discussion of the results of the computations, I would like to summarize briefly two important kinds of stability bifurcations which

[1] Although generally speaking this word is used in many physical situations in which there is a sudden change in properties for some critical parameter value.

typically occur in the study of spectral stability of Hamiltonian systems, and what each implies. When solving the eigenvalue problem, we will be looking for eigenvalues that are either purely imaginary (corresponding to a stable state) or have a non-zero real part (unstable). This is because, for some eigenvalue $\lambda = k + i\gamma$, we also have its complex conjugate $\bar{\lambda} = k - i\gamma$, its negative, $-\lambda = -k - i\gamma$, *and* its negative complex conjugate $-\bar{\lambda} = -k + i\gamma$, which is a result of the conservation of phase space volume (if we didn't have the negative, then there would be a direction in which everything shrinks but nowhere else is it growing, so it would not preserve volume). Therefore, if there is *any* non-zero real part of the eigenvalue, then there will always be an exponentially growing mode, which is unstable. Note that non-zero real part does not mean that the imaginary part must be zero! This will be important in a second. If we begin with a stable situation, then our eigenvalues are all imaginary. As we change the parameter, we approach the point of bifurcation, at which the system becomes unstable and we have some non-zero real parts in our eigenvalues. How this instability sets in determines the kind of bifurcation. It turns out that, generally speaking, there are two kinds of these transitions in Hamiltonian systems [24] [26]:

1. **Steady State Bifurcation** In this situation, a complex conjugate pair of eigenvalues "collide" at zero frequency, i.e. the origin of the complex plane, as you increase the parameter value up to the critical point, and then become *purely* real-valued after the critical point is passed. That is, they go from purely oscillatory to exponential growth and decay. See Figure 4.3 for a nice picture of what is happening on the complex plane. Intuitively, this happens, for example, in a "natural" (i.e. separable) Hamiltonian system when the potential energy surface changes concavity (so in one dimension it goes from an upward facing parabola, which is stable, to a downward facing one, which is unstable). One example of such a steady state bifurcation occurs in the symmetric case of two counter streaming ion beams with isothermal electrons [26]. Working with a system that goes through a steady state bifurcation allows us to use the Rayleigh-Ritz method for stability analysis [13], in which one uses trial functions to attempt to find a negative Rayleigh quotient, implying instability.
2. **Krein (or Hamiltonian-Hopf) Bifurcation** This second type of bifurcation occurs when *two* complex conjugate pairs, $\pm i\omega_1, \pm i\omega_2$ "collide" with the other (that is, $i\omega_1$ collides with $i\omega_2$) on the imaginary axis at some value of the parameter (the critical value) but *not* at the origin of the complex plane. When they collide they split off and, while they maintain a non-zero imaginary part, gain a non-zero real part, making them unstable. See Figure 4.4 for a nice picture of what I am talking about. One example of this kind of bifurcation can occur in the famous "Jeans instability," in which there are counter-flowing streams of self-gravitating

fluid past each other [26]. When these eigenvalues collide, it's typically due to the system having what is called a *negative-energy mode*, which are important because they cause instability when dissipation is added to the system [24].

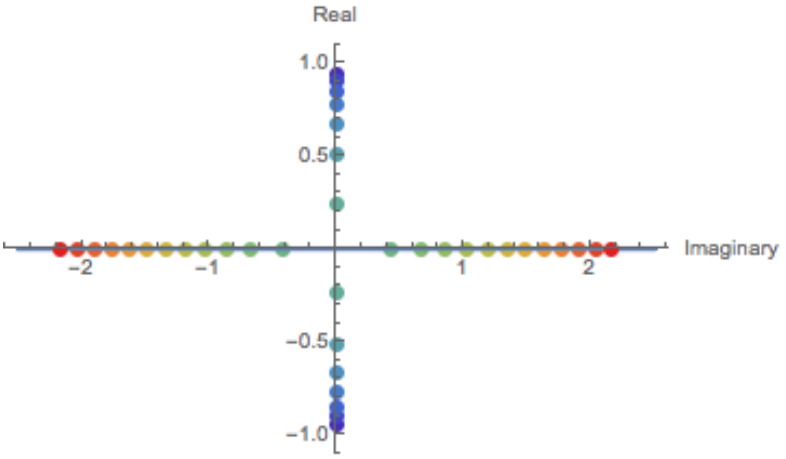


FIGURE 4.3: A typical example of a steady state bifurcation. The color of the dots represent the parameter values, increasing going from red to blue. Notice that at no point does the pair of eigenvalues become a mix of real and imaginary parts.

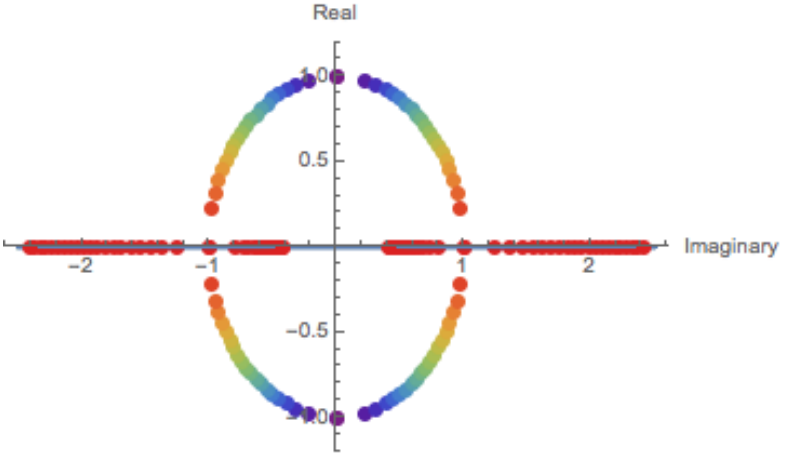


FIGURE 4.4: A typical example of a Krein or Hamiltonian-Hopf bifurcation. The color of the dots represent the parameter values, increasing going from red to blue. Here we are looking at two paris of eigenvalues colliding at i, and then going into a mix of real and imaginary.

For a more extensive review of these types of bifurcations, including a few examples, see Chapter 12 of *Nonlinear Physics Systems: Spectral Analysis, Stability and Bifurcation* [26].

4.3 Results

The resulting eigenvalues of interest of the matrix $\mathcal{H}^e$ are seen in Fig. 4.5. There is a single pair of eigenvalues that start out as purely real and that eventually become purely imaginary (those seen in the figure). This implies that our system is going through a *steady state bifurcation*. There are 10 other eigenvalues that start out and remain purely imaginary for all values of our parameter $\tilde{a}_3$. Finally, there are 6 eigenvalues that are exactly 0 for all values of our parameter $\tilde{a}_3$. These *neutral modes*, or *zero-frequency modes* arise from the fact that $\det J_{\alpha\beta} = 0$ due to the fact that we are working with a non-canonical system. Qualitatively, our results match those of Chandrasekhar, but as

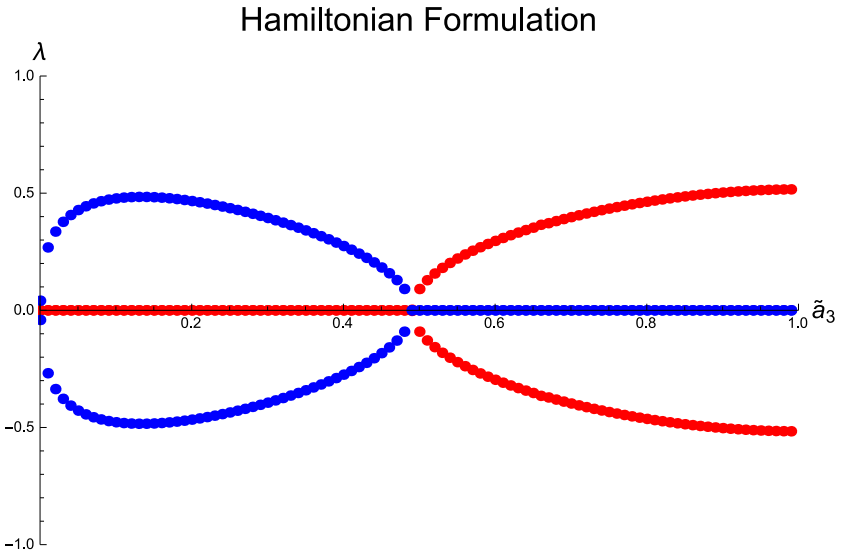


FIGURE 4.5: Here we are plotting the only pair of eigenvalues that start out having a non-zero real part as a function. The real part of the eigenvalues is in blue and the imaginary part in red. We are looking at a pair of eigenvalues undergoing a steady-state bifurcation. The rest of the eigenvalues are either all purely imaginary (and remain that way for all values of our parameter $\tilde{a}_3$) or are exactly zero. These results, however, do not seem to agree with Chandrasekhar's results seen in Fig. 4.6.

we will see below, there is a slight difference in the value of $\tilde{a}_3$ at which the bifurcation occurs.

I will briefly go over the procedure and results that Chandrasekhar obtained and then compare our results. Once again, the material from Chandrasekhar is taken from his 1965 paper [3], although it can also be found in Chapter 6 of his book [4]. Chandrasekhar begins first by manipulating the virial equation (Eq. 3.9) and expressing it in terms of $\vec{u}$, by using Eq. 3.2. His result is Eq. (18) from his paper:

$$\frac{d}{dt}\int_V \rho u_i x_j d\vec{x} = \int_V \rho u_i u_j d\vec{x} + \mathfrak{B}_{ij} + \Omega_3^2 (I_{ij} - \delta_{i3} I_{3i}) + 2\Omega_3 \epsilon_{il3} \int_V \rho u_l x_j d\vec{x} + \delta_{ij}\Pi. \quad (4.21)$$

One thing to note is that in order to get to this version of the virial equation one has to assume that Ω_3 is time-independent, which we weren't doing up to now, and we don't explicitly do in the Hamiltonian formulation, since everything is in terms of $\vec{U}$ and not $\vec{u}$, like it is here.[2] Chandrasekhar then perturbs around the equilibrium (we are still in the more general case where f and $\tilde{a}_3$ are unspecified) with a Lagrangian displacement $\vec{\xi}$, giving the fluid element at $\vec{x}$ a new position $\vec{x}+\vec{\xi}$ and a new velocity $\vec{u}+\Delta\vec{u}$, where $\Delta u_i = \partial\xi_i/\partial t + u_j(\partial\xi_i/\partial x_j)$. After then applying our linear velocity assumption, namely,

$$u_i = (K_{ij} + \Omega^*_{ij})x_j := Q_{ij}x_j,$$

the linearized virial equation then becomes:

$$\begin{aligned}&\frac{d^2V_{i;j}}{dt^2} - 2Q_{jl}\frac{dV_{i;l}}{dt} + Q_{jr}Q_{rl}V_{i;l} + Q_{is}Q_{sl}V_{j;l} = \\ &2\Omega_3\epsilon_{il3}\left(\frac{dV_{l;j}}{dt} - Q_{jk}V_{l;k} + Q_{lk}V_{j;k}\right) + \Omega_3^2\left(V_{ij} - \delta_{i3}V_{3j}\right) + \delta\mathfrak{B}_{ij} + \delta_{ij}\delta\Pi,\end{aligned} \tag{4.22}$$

where $V_{i;j} = \int_V \rho\xi_i x_j d\vec{x}$, and $V_{ij} = V_{i;j} + V_{j;i}$ (here we are following Chandrasekhar's notation). It is also important to note that $\vec{\xi}$ is solenoidal (i.e. divergence-free), meaning that

$$\frac{V_{11}}{a_1^2} + \frac{V_{22}}{a_2^2} + \frac{V_{33}}{a_3^2} = 0,$$

and that $\delta\mathfrak{B}_{ij}$ can be expressed in terms of V_{ij}, $\mathcal{A}_{ij}$, and B_{ij} (see equations (87) and (88) in [3]), where

$$\mathcal{A}_{ij} = \int_0^\infty \frac{du}{\Delta(u)(a_i^2+u)(a_j^2+u)}. \tag{4.23}$$

Finally, we will assume that

$$\vec{\xi}(\vec{x},t) = e^{\lambda t}\hat{\vec{\xi}}(\vec{x}),$$

and we have an eigenvalue problem on our hands. The equations become a bit messy and quite tedious, so I will skip the work and refer to Chandrasekhar's 1965 paper for the details. It turns out that all second-order virials[3] of Riemann Ellipsoids with $f \geq -2$ (that is, equilibrium ellipsoids for every possible value of $\tilde{a}_3$ and $f \geq -2$) are spectrally stable. However, for $f < -2$, which includes our value of f, this is not the case. In particular, for our specific value of $f = -(a_1^2+a_2^2)/(a_1a_2)$ an instability sets in at $\tilde{a}_2 = 0.29633$, $\tilde{a}_3 = 0.40733$ and is unstable for equilibrium values connecting this point and the origin of the $\tilde{a}_2$-$\tilde{a}_3$ graph (see Fig. 4.1) [3]. The characteristic equation

[2] I am not sure on whether or not this makes a difference and whether or not it is a reasonable assumption. Intuitively it doesn't make sense to me because if your axes are time dependent, then so is your angular momentum, which means that your angular velocity will be as well. But then again, it is *far* more likely that I am making a mistake in my reasoning... Anyways, I'm just thinking out loud.

[3] Note that this is not necessarily true for third-order virials (i.e. oscillations belonging to third harmonics).

for the modes that become unstable is

$$\lambda^4 + (4B_{13} + 4B_{23} + 2B_{12})\,\lambda^2 \qquad (4.24)$$
$$+ (4B_{13} - \Omega_3(K_{12} + \Omega_3))\,(4B_{23} + (\Omega_3(K_{21} - \Omega_3)) = 0.$$

Solving for λ, we get Figure 4.6. Note that our bifurcation point does *not* match the one that Chandrasekhar found, although qualitatively our results are very similar.

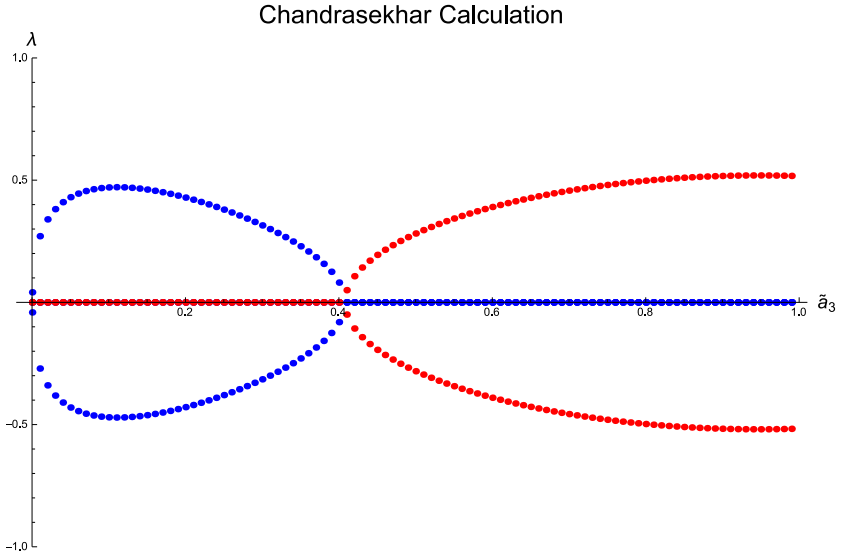


FIGURE 4.6: Here we are plotting λ as a function of $\tilde{a}_3$, with the real part of the eigenvalues in blue and the imaginary part in red. A clear steady-state bifurcation is seen at the value of $\tilde{a}_3 = 0.40733$, as Chandrasekhar claims.

Finally, it is worth noting that, in reference to Theorem 1.3, the Hessian matrix of F is indefinite, which unfortunately tells us nothing. The theorem which we are using is not necessary and sufficient, so having an indefinite Hessian matrix doesn't tell us anything about the instability of the system. This is, of course, assuming that our qualitative similarity to Chandrasekhar's results tells us that our physical results are right, but rather for incorrect parameter values.

All of the computations were done using *Mathematica*. See Appendix A for the *Mathematica* file in which the computations were made.

Chapter 5

Conclusion

5.1 Issues and Ideas

As one can see, figures 4.6 and 4.5 are not the same, even though they should be. There are many possible sources of error, including wrong implementation of the code, incorrect analytical expressions (a possible suspect is the calculation of $J_{\alpha\beta}$, since $\partial F/\partial z_\alpha = 0$ gives us the same equilibrium family that Chandrasekhar found, which seems to imply that this is correct), and maybe even numerical error, although I don't suspect this is true. What is strange is that, qualitatively, my results are very similar to those of Chandrasekhar (one unstable eigenvalue going through a steady-state bifurcation), so the physics *seems* right (or at least not far off) and there is nothing that is obviously wrong. I have re-calculated many of the analytical expressions more than once, gone through the code more than once, and done other troubleshooting things, all to no avail. When (if) I ever find the solution to this problem, I hope it will not be sitting under my nose... However, for the time being, let's assume that my results being qualitatively right implies that my results are physical, but it's just the parameter value that's different. This allows us to at least attempt to perform nonlinear stability analysis to our system which is what I did. Unfortunately, we saw that our Hessian matrix for F is indefinite, telling us nothing about the nonlinear stability of our system in the region of spectral stability. This leaves us with the same results (and uncertainties) that Chandrasekhar had and therefore does not directly resolve the issue between Riemann and Lebovitz/Chandrasekhar. However, since we are clearly going through a steady-state bifurcation, one possible resolution that I am looking into is the use of Rayleigh-Ritz method (using trial functions) and Lagrange's theorem in order to test for nonlinear instability in the spectrally stable regions.

There is also the withstanding issue of approaching the Rosensteel-Dirac bracket from the "other" direction, described in Section 2.2.4 of Chapter 2. This has turned out to be somewhat of a nightmare, as far as progress goes – every time I thought I had a breakthrough, it turned out that there was some issue further back in the calculations that I was not being very careful about. Don't get me wrong, I certainly learned (and am learning) a lot about the small subtleties around considering boundary terms, as well as dealing with Dirac delta functions – all of which are essential tools! The issue of relating this constrained fluid bracket to the Rosensteel-Dirac bracket is another problem, however, because there is no obvious way to transform the nasty integrals we have (and possibly boundary terms) to something proportional to $1/\mathrm{Tr}(\mathbf{I}^{-1})$. This is something else I will be working on in the near future.

5.2 Conclusion

In this thesis I presented the necessary mathematics background of differential topology and symplectic geometry to discuss Hamiltonian systems, and then presented a Hamiltonian formulation of the classical problem of studying the equilibria and stability of Riemann ellipsoids. I attempted to do an original derivation of the Rosensteel-Dirac bracket as well as study the stability of a certain (debated) family of equilibria using the Hamiltonian formulation for the first time. This project was started in the summer of 2014 at Woods Hole Oceanographic Institution's "Geophysical Fluid Dynamics" summer school in Woods Hole, Massachusetts. The work done in this thesis comprises of the work I did that summer and the next academic year (although the fall semester saw little work done due to graduate school applications and difficult classes). Although neither of the two attempts were successful, along the way I learned a great deal about Hamiltonian systems, fluid dynamics, stability theory, and, more specifically, Riemann ellipsoids. I consider this a great learning experience which I will take with me when I move on to graduate school.

Appendix A

Mathematica Code

A.1 Equilibrium Calculations

Below is the Mathematica Notebook in which I calculated the equilibrium values of the ellipsoid axes as well as Ω_3. Before we get to it, I will address the issue of non-dimensionalization which is implemented in the code (and more so in the stability calculations). Here are the following length, mass, and time scales that I will non-dimensionalize with from now on:

$$L = a_1 \tag{A.1}$$

$$M = \frac{4}{3}\pi\rho a_1^3 \tag{A.2}$$

$$T = (\pi\rho G)^{-1/2}\,. \tag{A.3}$$

Non-dimensional variables will appear with a tilde on top.

Riemann Ellipsoid Equilibrium Condition

```
a2 =.
a3 =.
```

```
Imat // MatrixForm
```

$$\begin{pmatrix} \frac{a1^2\, m}{5} & 0 & 0 \\ 0 & \frac{a2^2\, m}{5} & 0 \\ 0 & 0 & \frac{a3^2\, m}{5} \end{pmatrix}$$

```
Kmat2 = {{a1, 0, 0}, {0, a2, 0}, {0, 0, a3}} . {{0, Ω, 0}, {-Ω, 0, 0}, {0, 0, 0}} . {{1 / a1, 0, 0}, {0, 1 / a2, 0}, {0, 0, 1 / a3}} - {{0, Ω, 0}, {-Ω, 0, 0}, {0, 0, 0}};
```

```
Om = {{0, Ω, 0}, {-Ω, 0, 0}, {0, 0, 0}};
```

```
Kmat2 // MatrixForm
```

$$\begin{pmatrix} 0 & -\Omega + \frac{a1\,\Omega}{a2} & 0 \\ \Omega - \frac{a2\,\Omega}{a1} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

```
Nmat2 = Imat.Transpose[Kmat2];
```

```
Inverse[Imat] // MatrixForm
```

$$\begin{pmatrix} \frac{5}{a1^2\, m} & 0 & 0 \\ 0 & \frac{5}{a2^2\, m} & 0 \\ 0 & 0 & \frac{5}{a3^2\, m} \end{pmatrix}$$

Since we don't want to integrate the gravitational part yet, we will use a dummy matrix for that.

```
V0equil = Table[0, {a, 1, 3}, {d, 1, 3}];
```

```
Wmattest = ReplacePart[V0equil, {{1, 1} → W11, {2, 2} → W22, {3, 3} → W33}];
```

```
Wmattest // MatrixForm
```

$$\begin{pmatrix} W11 & 0 & 0 \\ 0 & W22 & 0 \\ 0 & 0 & W33 \end{pmatrix}$$

```
z0 = Table[0., {a, 1, 18}];

c1 =.
c2 =.
c3 =.

Fequil[c1_, c2_, c3_] := ReplacePart[z0,
    Flatten[Join[Table[{j + 3 * (i - 1)} → -1/2 * (Transpose[Inverse[Imat]].
            Nmat2.Transpose[Nmat2].Transpose[Inverse[Imat]])[[i, j]] +
          Wmattest[[i, j]] + c1 * Inverse[Imat][[i, j]] - c2 *
           (Transpose[Inverse[Imat]].Transpose[Nmat2].Transpose[Inverse[Imat]])[[
            i, j]] - c3 * (Transpose[Inverse[Imat]].Nmat2.Transpose[Imat].
            Transpose[Nmat2].Transpose[Inverse[Imat]])[[i, j]] +
          c3 * (Transpose[Nmat2].Transpose[Inverse[Imat]].Nmat2)[[i, j]],
        {i, 1, 3}, {j, 1, 3}], Table[{j + 3 * (i - 1) + 9} →
         (Inverse[Imat].Nmat2)[[i, j]] - Om[[i, j]] +
          c2 * Inverse[Imat][[i, j]] + c3 * (Inverse[Imat].Nmat2.Imat)[[i, j]] +
          c3 * (Transpose[Inverse[Imat]].Nmat2.Transpose[Imat])[[i, j]] -
          2 * c3 * Nmat2[[j, i]], {i, 1, 3}, {j, 1, 3}]]];

Chop[Fequil[c1, c2, c3], 10^-10]
```

$$\left\{\frac{5\,c1}{a1^2\,m} + W11 + \frac{1}{5}\,a2^2\,c3\,m\left(-\Omega + \frac{a1\,\Omega}{a2}\right)^2 - \frac{1}{2}\left(\Omega - \frac{a2\,\Omega}{a1}\right)^2 - \frac{1}{5}\,a2^2\,c3\,m\left(\Omega - \frac{a2\,\Omega}{a1}\right)^2,\right.$$
$$-\frac{5\,c2\left(-\Omega + \frac{a1\,\Omega}{a2}\right)}{a1^2\,m}, 0, -\frac{5\,c2\left(\Omega - \frac{a2\,\Omega}{a1}\right)}{a2^2\,m},$$
$$\frac{5\,c1}{a2^2\,m} + W22 - \frac{1}{2}\left(-\Omega + \frac{a1\,\Omega}{a2}\right)^2 - \frac{1}{5}\,a1^2\,c3\,m\left(-\Omega + \frac{a1\,\Omega}{a2}\right)^2 + \frac{1}{5}\,a1^2\,c3\,m\left(\Omega - \frac{a2\,\Omega}{a1}\right)^2, 0,$$
$$0, 0, \frac{5\,c1}{a3^2\,m} + W33, \frac{5\,c2}{a1^2\,m}, -\frac{a2\,\Omega}{a1} - \frac{2}{5}\,a2^2\,c3\,m\left(-\Omega + \frac{a1\,\Omega}{a2}\right) + \frac{2}{5}\,a2^2\,c3\,m\left(\Omega - \frac{a2\,\Omega}{a1}\right),$$
$$\left.0, \frac{a1\,\Omega}{a2} + \frac{2}{5}\,a1^2\,c3\,m\left(-\Omega + \frac{a1\,\Omega}{a2}\right) - \frac{2}{5}\,a1^2\,c3\,m\left(\Omega - \frac{a2\,\Omega}{a1}\right), \frac{5\,c2}{a2^2\,m}, 0, 0, 0, \frac{5\,c2}{a3^2\,m}\right\}$$

```
Solve[5 c2/a3^2 m == 0, c2]
```

$$\{\{c2 \to 0\}\}$$

```
c2 = 0;
```

$$\text{FullSimplify}\left[\left\{\frac{5\,c1}{a1^2\,m}+W11+\frac{1}{5}a2^2\,c3\,m\left(-\Omega+\frac{a1\,\Omega}{a2}\right)^2-\frac{1}{2}\left(\Omega-\frac{a2\,\Omega}{a1}\right)^2-\frac{1}{5}a2^2\,c3\,m\left(\Omega-\frac{a2\,\Omega}{a1}\right)^2,\right.\right.$$
$$-\frac{5\,c2\left(-\Omega+\frac{a1\,\Omega}{a2}\right)}{a1^2\,m},\ 0,\ -\frac{5\,c2\left(\Omega-\frac{a2\,\Omega}{a1}\right)}{a2^2\,m},$$
$$\frac{5\,c1}{a2^2\,m}+W22-\frac{1}{2}\left(-\Omega+\frac{a1\,\Omega}{a2}\right)^2-\frac{1}{5}a1^2\,c3\,m\left(-\Omega+\frac{a1\,\Omega}{a2}\right)^2+\frac{1}{5}a1^2\,c3\,m\left(\Omega-\frac{a2\,\Omega}{a1}\right)^2,\ 0,$$
$$0,\ 0,\ \frac{5\,c1}{a3^2\,m}+W33,\ \frac{5\,c2}{a1^2\,m},\ -\frac{a2\,\Omega}{a1}-\frac{2}{5}a2^2\,c3\,m\left(-\Omega+\frac{a1\,\Omega}{a2}\right)+\frac{2}{5}a2^2\,c3\,m\left(\Omega-\frac{a2\,\Omega}{a1}\right),$$
$$\left.\left.0,\ \frac{a1\,\Omega}{a2}+\frac{2}{5}a1^2\,c3\,m\left(-\Omega+\frac{a1\,\Omega}{a2}\right)-\frac{2}{5}a1^2\,c3\,m\left(\Omega-\frac{a2\,\Omega}{a1}\right),\ \frac{5\,c2}{a2^2\,m},\ 0,\ 0,\ 0,\ \frac{5\,c2}{a3^2\,m}\right\}\right]$$

$$\left\{\frac{1}{10\,a1^2\,m}\left(50\,c1+10\,a1^2\,m\,W11+(a1-a2)^2\,m\,(-5+2\,(a1-a2)\,(a1+a2)\,c3\,m)\,\Omega^2\right),\ 0,\ 0,\right.$$
$$0,\ \frac{1}{10\,a2^2\,m}\left(50\,c1+10\,a2^2\,m\,W22+(a1-a2)^2\,m\left(-5+2\left(-a1^2+a2^2\right)c3\,m\right)\Omega^2\right),\ 0,\ 0,\ 0,$$
$$\left.\frac{5\,c1}{a3^2\,m}+W33,\ 0,\ -\frac{a2\left(5+2\,(a1-a2)^2\,c3\,m\right)\Omega}{5\,a1},\ 0,\ \frac{a1\left(5+2\,(a1-a2)^2\,c3\,m\right)\Omega}{5\,a2},\ 0,\ 0,\ 0,\ 0,\ 0\right\}$$

$$\text{Solve}\left[-\frac{a2\left(5+2\,(a1-a2)^2\,c3\,m\right)\Omega}{5\,a1}==0,\ c3\right]$$

$$\left\{\left\{c3\to-\frac{5}{2\,(a1-a2)^2\,m}\right\}\right\}$$

$$c3=-\frac{5}{2\,(a1-a2)^2\,m};$$

$$\text{FullSimplify}\left[\left\{\frac{5\,c1}{a1^2\,m}+W11+\frac{1}{5}a2^2\,c3\,m\left(-\Omega+\frac{a1\,\Omega}{a2}\right)^2-\frac{1}{2}\left(\Omega-\frac{a2\,\Omega}{a1}\right)^2-\frac{1}{5}a2^2\,c3\,m\left(\Omega-\frac{a2\,\Omega}{a1}\right)^2,\right.\right.$$
$$-\frac{5\,c2\left(-\Omega+\frac{a1\,\Omega}{a2}\right)}{a1^2\,m},\ 0,\ -\frac{5\,c2\left(\Omega-\frac{a2\,\Omega}{a1}\right)}{a2^2\,m},$$
$$\frac{5\,c1}{a2^2\,m}+W22-\frac{1}{2}\left(-\Omega+\frac{a1\,\Omega}{a2}\right)^2-\frac{1}{5}a1^2\,c3\,m\left(-\Omega+\frac{a1\,\Omega}{a2}\right)^2+\frac{1}{5}a1^2\,c3\,m\left(\Omega-\frac{a2\,\Omega}{a1}\right)^2,\ 0,$$
$$0,\ 0,\ \frac{5\,c1}{a3^2\,m}+W33,\ \frac{5\,c2}{a1^2\,m},\ -\frac{a2\,\Omega}{a1}-\frac{2}{5}a2^2\,c3\,m\left(-\Omega+\frac{a1\,\Omega}{a2}\right)+\frac{2}{5}a2^2\,c3\,m\left(\Omega-\frac{a2\,\Omega}{a1}\right),$$
$$\left.\left.0,\ \frac{a1\,\Omega}{a2}+\frac{2}{5}a1^2\,c3\,m\left(-\Omega+\frac{a1\,\Omega}{a2}\right)-\frac{2}{5}a1^2\,c3\,m\left(\Omega-\frac{a2\,\Omega}{a1}\right),\ \frac{5\,c2}{a2^2\,m},\ 0,\ 0,\ 0,\ \frac{5\,c2}{a3^2\,m}\right\}\right]$$

$$\left\{\frac{5\,c1}{a1^2\,m}+W11-\Omega^2+\frac{a2\,\Omega^2}{a1},\ 0,\ 0,\ 0,\ \frac{5\,c1}{a2^2\,m}+W22-\Omega^2+\frac{a1\,\Omega^2}{a2},\right.$$
$$\left.0,\ 0,\ 0,\ \frac{5\,c1}{a3^2\,m}+W33,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0\right\}$$

```
Solve[5 c1/(a3^2 m) + W33 == 0, c1]
```

$$\left\{\left\{c1 \to -\frac{1}{5}\, a3^2\, m\, W33\right\}\right\}$$

```
c1 = -1/5 a3^2 m W33;
```

```
FullSimplify[{5 c1/(a1^2 m) + W11 + 1/5 a2^2 c3 m (-Ω + a1 Ω/a2)^2 - 1/2 (Ω - a2 Ω/a1)^2 - 1/5 a2^2 c3 m (Ω - a2 Ω/a1)^2,
  -(5 c2 (-Ω + a1 Ω/a2))/(a1^2 m), 0, -(5 c2 (Ω - a2 Ω/a1))/(a2^2 m),
  5 c1/(a2^2 m) + W22 - 1/2 (-Ω + a1 Ω/a2)^2 - 1/5 a1^2 c3 m (-Ω + a1 Ω/a2)^2 + 1/5 a1^2 c3 m (Ω - a2 Ω/a1)^2, 0,
  0, 0, 5 c1/(a3^2 m) + W33, 5 c2/(a1^2 m), -(a2 Ω/a1) - 2/5 a2^2 c3 m (-Ω + a1 Ω/a2) + 2/5 a2^2 c3 m (Ω - a2 Ω/a1),
  0, a1 Ω/a2 + 2/5 a1^2 c3 m (-Ω + a1 Ω/a2) - 2/5 a1^2 c3 m (Ω - a2 Ω/a1), 5 c2/(a2^2 m), 0, 0, 0, 5 c2/(a3^2 m)}]
```

$$\left\{W11 + \frac{-a3^2\, W33 + a1\,(-a1 + a2)\,\Omega^2}{a1^2},\ 0,\ 0,\ 0,\right.$$
$$\left. W22 + \frac{-a3^2\, W33 + (a1 - a2)\, a2\, \Omega^2}{a2^2},\ 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\right\}$$

```
Solve[W22 + (-a3^2 W33 + (a1 - a2) a2 Ω^2)/a2^2 == 0, Ω]
```

$$\left\{\left\{\Omega \to -\frac{\sqrt{-a2^2\, W22 + a3^2\, W33}}{\sqrt{a1\, a2 - a2^2}}\right\}, \left\{\Omega \to \frac{\sqrt{-a2^2\, W22 + a3^2\, W33}}{\sqrt{a1\, a2 - a2^2}}\right\}\right\}$$

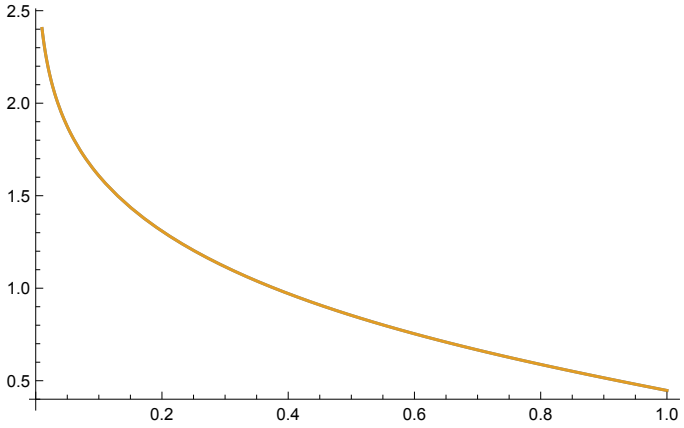
```
Ω = Sqrt[-a2^2 W22 + a3^2 W33]/Sqrt[a1 a2 - a2^2];
```

```
FullSimplify[Ω^2]
```

$$\frac{-a2^2\, W22 + a3^2\, W33}{(a1 - a2)\, a2}$$

```
Ωtest[a2_, a3_] := Sqrt[(-a2^2 dA2c[a2, a3] + a3^2 dA3c[a2, a3])/((1 - a2) a2)]
```

```
Plot[{Ω3[FFine[a3], a3], Ωtest[FFine[a3], a3]}, {a3, .01, 1}]
```



So this actually gives the same Omega that Chandra calculated! Now let's solve the remaining equation and get a function of the a_3 axis in terms of a_2.

```
FullSimplify[{5 c1/(a1^2 m) + W11 + 1/5 a2^2 c3 m (-Ω + a1 Ω/a2)^2 - 1/2 (Ω - a2 Ω/a1)^2 - 1/5 a2^2 c3 m (Ω - a2 Ω/a1)^2,
 -(5 c2 (-Ω + a1 Ω/a2))/(a1^2 m), 0, -(5 c2 (Ω - a2 Ω/a1))/(a2^2 m),
 5 c1/(a2^2 m) + W22 - 1/2 (-Ω + a1 Ω/a2)^2 - 1/5 a1^2 c3 m (-Ω + a1 Ω/a2)^2 + 1/5 a1^2 c3 m (Ω - a2 Ω/a1)^2, 0,
 0, 0, 5 c1/(a3^2 m) + W33, 5 c2/(a1^2 m), -(a2 Ω/a1) - 2/5 a2^2 c3 m (-Ω + a1 Ω/a2) + 2/5 a2^2 c3 m (Ω - a2 Ω/a1),
 0, a1 Ω/a2 + 2/5 a1^2 c3 m (-Ω + a1 Ω/a2) - 2/5 a1^2 c3 m (Ω - a2 Ω/a1), 5 c2/(a2^2 m), 0, 0, 0, 5 c2/(a3^2 m)}]
```

$$\left\{\frac{a1\ a2\ (a1\ W11 + a2\ W22) - (a1 + a2)\ a3^2\ W33}{a1^2\ a2}, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\right\}$$

```
Expand[(a1 a2 (a1 W11 + a2 W22) - (a1 + a2) a3^2 W33)/(a1^2 a2)]
```

$$W11 + \frac{a2\ W22}{a1} - \frac{a3^2\ W33}{a1^2} - \frac{a3^2\ W33}{a1\ a2}$$

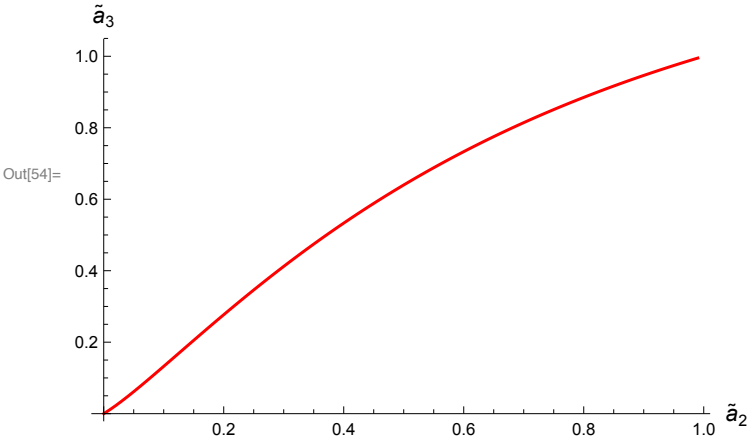
```
dA1[a3_?NumericQ] :=
  3/4 NIntegrate[1/(Sqrt[(1+u) (a2^2+u) (a3^2+u)] * (1+u)), {u, 0, Infinity}];
dA2[a3_?NumericQ] :=
  3/4 NIntegrate[1/(Sqrt[(1+u) (a2^2+u) (a3^2+u)] * (a2^2+u)), {u, 0, Infinity}];
dA3[a3_?NumericQ] :=
  3/4 NIntegrate[1/(Sqrt[(1+u) (a2^2+u) (a3^2+u)] * (a3^2+u)), {u, 0, Infinity}];
```

$$W11 + \frac{a2\,W22}{a1} - \frac{a3^2\,W33}{a1^2} - \frac{a3^2\,W33}{a1\,a2}$$

```
AbsoluteTiming[
 a3valsFequil = Table[a3 /. Assuming[a3 > 0, FindRoot[dA1[a3] + a2 dA2[a3] -
      a3^2 dA3[a3] - a3^2 dA3[a3]/a2 == 0, {a3, 0, 1}]], {a2, 0.0001, 1., .01}];]
datatestFequil = Partition[Riffle[Table[a2, {a2, 0.0001, 1., .01}], a3valsFequil], 2];
```

In[54]:=
```
ListPlot[{datatestFequil}, Joined → True,
 AxesLabel → {Style["ã₂", FontSize → 12], Style["ã₃", FontSize → 12]},
 PlotLabel → Style["Equilibria , f = - (a₁^2 + a₂^2)/(a₁ a₂)", FontSize → 15], PlotStyle → Red]
```

Out[54]=



Chandrasekhar Equilibrium Conditions

First we begin by defining the same old equilibrium conditions:

I am working through the equilibrium condition for $x = -1$. In this case, the equilibrium condition is the following (Note: $\tilde{a_2} = a_2/a_1$ and $\tilde{a_3} = a_3/a_1$)

$$\tilde{a}_2\left(\tilde{A}_1 - \tilde{A}_2 + \frac{1-\tilde{a}_2^2}{\tilde{a}_2^2}\,\tilde{a}_3^2\,\tilde{A}_3\right) = \left(\tilde{A}_1 - \tilde{A}_2\,\tilde{a}_2^2\right)$$

where

$$\tilde{A}_i = \int_0^\infty \frac{d\,\tilde{u}}{\Delta\left(\tilde{a}_i^2+\tilde{u}\right)}, \qquad \Delta = \sqrt{(1+\tilde{u})\left(\tilde{a}_2^2+\tilde{u}\right)\left(\tilde{a}_3^2+\tilde{u}\right)}\,.$$

Note that everything here is dimensionless!

```
A1[a3_?NumericQ] :=
  NIntegrate[1/(Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (1 + u)), {u, 0, Infinity}, AccuracyGoal → 8];
A2[a3_?NumericQ] := NIntegrate[1/(Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (a2^2 + u)),
   {u, 0, Infinity}, AccuracyGoal → 8];
A3[a3_?NumericQ] := NIntegrate[1/(Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (a3^2 + u)),
   {u, 0, Infinity}, AccuracyGoal → 8];
```

Below is the calculation which tells us $\tilde{a}_2(\tilde{a}_3)$ as a function of $\tilde{a}_3$, which is our parameter, such that we are in an equilibrium state along the $x = -1$ branch.

```
AbsoluteTiming[a3vals =
   Table[a3 /. Assuming[a3 > 0, FindRoot[a2 * (A1[a3] - A2[a3] + (1 - a2^2)/a2^2 a3^2 A3[a3]) -
         (A1[a3] - A2[a3] * a2^2) == 0, {a3, 0, 1}]], {a2, 0.0001, 1., .01}];]
```

```
AbsoluteTiming[a3valsfine =
  Table[a3 /. Assuming[a3 > 0, FindRoot[a2 * ((A1[a3] - A2[a3] + (1 - a2^2)/a2^2 a3^2 A3[a3]) -
      (A1[a3] - A2[a3] * a2^2) == 0, {a3, 0, 1}]], {a2, 0.0001, 1., .001}];]
a3valsfine
```

Click on “a3vals” and “a3valsfine” to see the data and not have to re-compute it!

Now, we want to define $\tilde{a}_2$ as a function of our parameter $\tilde{a}_3$, so we use interpolation:

In[13]:=
```
interpdata = Partition[Riffle[a3valsfine, Table[a2, {a2, 0.0001, 1., .001}]], 2];
```

In[14]:=
```
FFine = Interpolation[interpdata]
```

Out[14]= InterpolatingFunction[Domain : {{0.0001 , 1.}} Output : scalar]

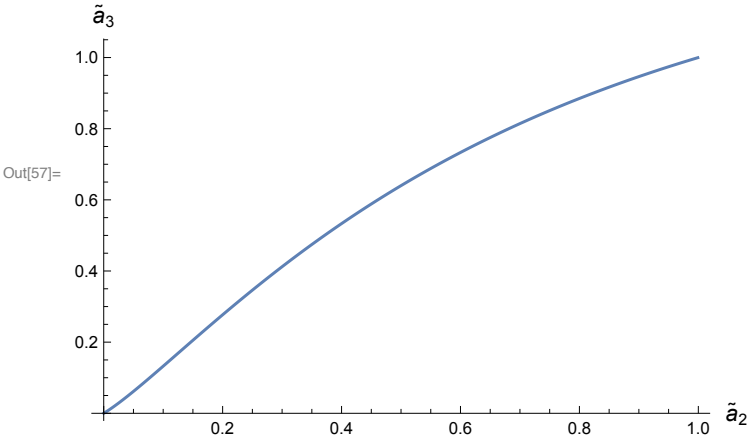
```
FFine[.56]
```

```
0.423889
```

In[35]:=
```
chandraplot = Partition[Riffle[Table[a2, {a2, 0.0001, 1., .001}], a3valsfine], 2];
```

In[57]:=
```
ListPlot[chandraplot, Joined → True,
 AxesLabel → {Style["ã₂", FontSize → 12], Style["ã₃", FontSize → 12]},
 PlotLabel → Style["Chandrasekhar Equilibria , f = - (a₁² + a₂²)/(a₁ a₂)", FontSize → 15]]
```

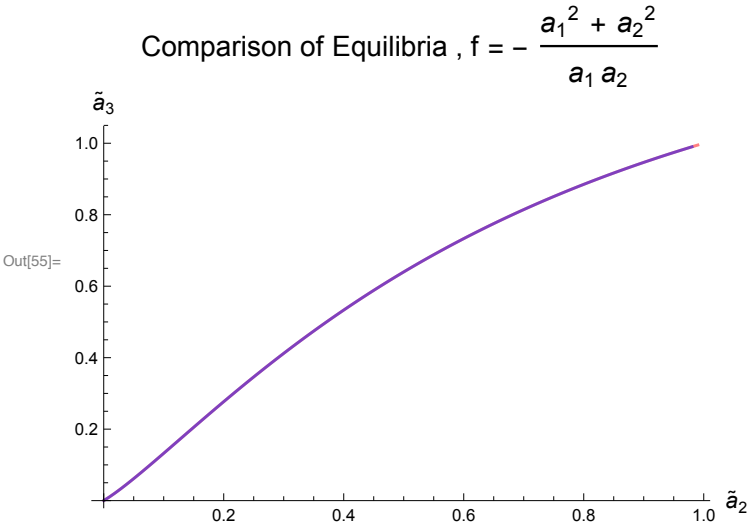
Out[57]=



Comparison

In[55]:=
```
ListPlot[{datatestFequil, Table[{FFine[a3], a3}, {a3, 0.0001, 1, .01}]},
 Joined → True, AxesLabel → {Style["ã2", FontSize → 12], Style["ã3", FontSize → 12]},
 PlotLabel → Style["Comparison of Equilibria , f = - (a1^2 + a2^2)/(a1 a2)", FontSize → 15],
 PlotStyle → {Directive[Red, Opacity[.5]], Directive[Blue, Opacity[.5]]}]
```

Out[55]=



A.2 Stability Calculations

Below is the Mathematica Notebook in which I calculated the spectral stability of the ellipsoids.

Riemann Ellipsoid Stability Calculations

Calculating J

So let's begin calculating J

```
Ii[a2_, a3_] := 0.2 * a2 * a3 ({{1, 0, 0}, {0, a2^2, 0}, {0, 0, a3^2}});

invIi[a2_, a3_] := 5.0/(a2 * a3) * ({{1, 0, 0}, {0, 1/a2^2, 0}, {0, 0, 1/a3^2}});

B12[a2_, a3_] := NIntegrate[a2 * a3 * u/(Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (1 + u) * (a2^2 + u)), {u, 0, Infinity}, AccuracyGoal → 8];

Ω3[a2_, a3_] := Sqrt[B12[a2, a3]];

Ω[a2_, a3_] := ({{0, Ω3[a2, a3], 0}, {-Ω3[a2, a3], 0, 0}, {0, 0, 0}});

({{a1, 0, 0}, {0, a2, 0}, {0, 0, a3}}) . ({{0, Ω, 0}, {-Ω, 0, 0}, {0, 0, 0}}) . ({{1/a1, 0, 0}, {0, 1/a2, 0}, {0, 0, 1/a3}}) - ({{0, Ω, 0}, {-Ω, 0, 0}, {0, 0, 0}}) // MatrixForm
```

$$\begin{pmatrix} 0 & -\Omega + \frac{a1\,\Omega}{a2} & 0 \\ \Omega - \frac{a2\,\Omega}{a1} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

```
({{1, 0, 0}, {0, a2, 0}, {0, 0, a3}}) . ({{0, Ω, 0}, {-Ω, 0, 0}, {0, 0, 0}}) . ({{1, 0, 0}, {0, 1/a2, 0}, {0, 0, 1/a3}}) - ({{0, Ω, 0}, {-Ω, 0, 0}, {0, 0, 0}}) // MatrixForm
```

$$\begin{pmatrix} 0 & -\Omega + \frac{\Omega}{a2} & 0 \\ \Omega - a2\,\Omega & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

```
Kt[a2_, a3_] := {{0, -Ω3[a2, a3] + Ω3[a2,a3]/a2, 0},
                 {Ω3[a2, a3] - a2 Ω3[a2, a3], 0, 0},
                 {0, 0, 0}};

Ni[a2_, a3_] := Ii[a2, a3].Transpose[Kt[a2, a3]]

J0[a2_, a3_] := Table[0., {a, 1, 18}, {b, 1, 18}];

J[a2_, a3_] :=
  ReplacePart[J0[a2, a3], Flatten[Join[Table[{i, j} → 0., {i, 1, 9}, {j, 1, 9}],
     Table[{j + 3 * (i - 1), l + 3 (k - 1) + 9} → IdentityMatrix[3][[i, l]]
          Ii[a2, a3][[j, k]] + IdentityMatrix[3][[j, l]] Ii[a2, a3][[i, k]] -
        (2 IdentityMatrix[3][[i, j]] IdentityMatrix[3][[k, l]]) / Tr[invIi[a2, a3]],
      {i, 1, 3}, {j, 1, 3}, {k, 1, 3}, {l, 1, 3}], Table[
      {j + 3 * (i - 1) + 9, l + 3 (k - 1)} → -IdentityMatrix[3][[k, j]] Ii[a2, a3][[l, i]] -
        IdentityMatrix[3][[l, j]] Ii[a2, a3][[k, i]] +
        (2 IdentityMatrix[3][[i, j]] IdentityMatrix[3][[k, l]]) / Tr[invIi[a2, a3]],
      {i, 1, 3}, {j, 1, 3}, {k, 1, 3}, {l, 1, 3}], Table[
      {j + 3 * (i - 1) + 9, l + 3 (k - 1) + 9} → IdentityMatrix[3][[i, l]] Ni[a2, a3][[k, j]] -
        IdentityMatrix[3][[j, k]] Ni[a2, a3][[i, l]] + 1/Tr[invIi[a2, a3]]
         (IdentityMatrix[3][[i, j]] ((invIi[a2, a3].Ni[a2, a3])[[k, l]] +
              (invIi[a2, a3].Ni[a2, a3])[[l, k]]) - IdentityMatrix[3][[k, l]]
            ((invIi[a2, a3].Ni[a2, a3])[[i, j]] + (invIi[a2, a3].Ni[a2, a3])[[
               j, i]])), {i, 1, 3}, {j, 1, 3}, {k, 1, 3}, {l, 1, 3}]]];

FFine[.56]

0.423889

AbsoluteTiming[jtest = J[FFine[.56], .56];]

{22.645509, Null}
```

```
jtest + Transpose[jtest] // MatrixForm
```

```
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
```

Okay, so J is in fact anti-symmetric.

Confirming Equilibrium Conditions

```
dA1c[a2_, a3_] :=
  NIntegrate[(a2 * a3) / (Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (1 + u)), {u, 0, Infinity}, AccuracyGoal → 8];

dA2c[a2_, a3_] := NIntegrate[(a2 * a3) / (Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (a2^2 + u)),
   {u, 0, Infinity}, AccuracyGoal → 8];

dA3c[a2_, a3_] := NIntegrate[(a2 * a3) / (Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (a3^2 + u)),
   {u, 0, Infinity}, AccuracyGoal → 8];

w0 = Table[0, {a, 1, 3}, {b, 1, 3}];

Wfirstdev[a2_, a3_] := ReplacePart[w0,
   {{1, 1} → dA1c[a2, a3], {2, 2} → dA2c[a2, a3], {3, 3} → dA3c[a2, a3]}];

z0 = Table[0., {a, 1, 18}];
```

Let's also define the coefficients now, since they won't change.

```
c1[a2_, a3_] := -1/5 a2 * a3^3 dA3c[a2, a3]

c3[a2_, a3_] := -5/(2 * a2 * a3 * (1 - a2)^2)
```

Okay, now we are ready to calculate the gradient (see handwritten notes).

```
Fequil[a2_, a3_] :=
  ReplacePart[z0, Flatten[Join[Table[{j + 3 * (i - 1)} → -1/2 * (Transpose[invIi[a2, a3]].
            Ni[a2, a3].Transpose[Ni[a2, a3]].Transpose[invIi[a2, a3]])[[i, j]] +
        Wfirstdev[a2, a3][[i, j]] + c1[a2, a3] * invIi[a2, a3][[i, j]] -
        c3[a2, a3] * (Transpose[invIi[a2, a3]].Ni[a2, a3].Transpose[Ii[a2, a3]].
            Transpose[Ni[a2, a3]].Transpose[invIi[a2, a3]])[[i, j]] + c3[a2, a3] *
         (Transpose[Ni[a2, a3]].Transpose[invIi[a2, a3]].Ni[a2, a3])[[i, j]],
      {i, 1, 3}, {j, 1, 3}], Table[{j + 3 * (i - 1) + 9} →
       (invIi[a2, a3].Ni[a2, a3])[[i, j]] - Ω[a2, a3][[i, j]] +
        c3[a2, a3] * (invIi[a2, a3].Ni[a2, a3].Ii[a2, a3])[[i, j]] + c3[a2, a3] *
         (Transpose[invIi[a2, a3]].Ni[a2, a3].Transpose[Ii[a2, a3]])[[i, j]] -
        2 * c3[a2, a3] * Ni[a2, a3][[j, i]], {i, 1, 3}, {j, 1, 3}]]];

Dimensions[Fequil[FFine[.56], .56]]
{18}
```

Okay, so we have defined the gradient of F. Let's see if, for all values of a_3 and $a_2(a_3)$, we get that it is the zero vector (just to confirm once again that we are doing things right).

```
AbsoluteTiming[Chop[Fequil[FFine[.57], .57], 10^-10]]
{3.690962, {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0}}

AbsoluteTiming[Chop[Fequil[FFine[.6], .6], 10^-10]]
{3.733727, {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0}}

AbsoluteTiming[
 ListPlot[Table[{a3, Norm[Fequil[FFine[.56], .56]]}, {a3, 0.0001, 1., .01}]]]
$Aborted
```

Yay, it worked! Let's move on.

Now let's go calculate the Hessian matrix (L)

```
A11[a2_, a3_] := NIntegrate[
  1/(Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (1 + u) * (1 + u)), {u, 0, Infinity}, AccuracyGoal → 8];

A12[a2_, a3_] := NIntegrate[1/(Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (1 + u) * (a2^2 + u)),
  {u, 0, Infinity}, AccuracyGoal → 8];

A13[a2_, a3_] := NIntegrate[1/(Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (1 + u) * (a3^2 + u)),
  {u, 0, Infinity}, AccuracyGoal → 8];

A22[a2_, a3_] := NIntegrate[1/(Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (a2^2 + u) * (a2^2 + u)),
  {u, 0, Infinity}, AccuracyGoal → 8];

A23[a2_, a3_] := NIntegrate[1/(Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (a3^2 + u) * (a2^2 + u)),
  {u, 0, Infinity}, AccuracyGoal → 8];

A33[a2_, a3_] := NIntegrate[1/(Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (a3^2 + u) * (a3^2 + u)),
  {u, 0, Infinity}, AccuracyGoal → 8];

V0[a2_, a3_] := Table[0, {a, 1, 3}, {b, 1, 3}, {c, 1, 3}, {d, 1, 3}];

W[a2_, a3_] := ReplacePart[V0[a2, a3], {{1, 1, 1, 1} → -45/8 A11[a2, a3],
    {1, 1, 2, 2} → -15/8 A12[a2, a3], {1, 1, 3, 3} → -15/8 A13[a2, a3],
    {2, 2, 1, 1} → -15/8 A12[a2, a3], {2, 2, 2, 2} → -45/8 * A22[a2, a3],
    {2, 2, 3, 3} → -15/8 A23[a2, a3], {3, 3, 1, 1} → -15/8 A13[a2, a3],
    {3, 3, 2, 2} → -15/8 A23[a2, a3], {3, 3, 3, 3} → -45/8 A33[a2, a3]}];
```

And now we can finally add it all together to make L:

```
L0[a2_, a3_] := Table[0., {a, 1, 18}, {b, 1, 18}];

L[a2_, a3_] :=
  ReplacePart[L0[a2, a3], Flatten[Join[Table[{b + 3 * (a - 1), j + 3 * (i - 1)} →
       .5 * ((Transpose[invIi[a2, a3]].Ni[a2, a3].Transpose[Ni[a2, a3]].
             Transpose[invIi[a2, a3]])[[a, j]]) * (invIi[a2, a3][[b, i]]) +
        .5 * (invIi[a2, a3][[j, a]]) * ((invIi[a2, a3].Ni[a2, a3].
             Transpose[Ni[a2, a3]].invIi[a2, a3])[[b, i]]) + W[a2, a3][[a, b, i, j]] -
        c1[a2, a3] * (invIi[a2, a3][[i, a]]) * (invIi[a2, a3][[b, j]]) +
        c3[a2, a3] * (((invIi[a2, a3].Ni[a2, a3].Ii[a2, a3].Transpose[Ni[a2, a3]].
                invIi[a2, a3])[[b, i]]) * ((invIi[a2, a3])[[j, a]]) +
           ((invIi[a2, a3].Ni[a2, a3].Ii[a2, a3].Transpose[Ni[a2, a3]].
                invIi[a2, a3])[[j, a]]) * (invIi[a2, a3][[b, i]]) -
           ((invIi[a2, a3].Ni[a2, a3])[[j, a]]) * ((Transpose[invIi[a2, a3]].
                Ni[a2, a3])[[i, b]]) - ((invIi[a2, a3].Ni[a2, a3])[[b, i]]) *
            ((Transpose[invIi[a2, a3]].Ni[a2, a3])[[a, j]])), {i, 1, 3},
      {j, 1, 3}, {a, 1, 3}, {b, 1, 3}], Table[{b + 3 * (a - 1), j + 3 * (i - 1) + 9} →
       -.5 * ((invIi[a2, a3])[[i, a]]) * ((invIi[a2, a3].Ni[a2, a3])[[b, j]]) -
        .5 * (((invIi[a2, a3])[[b, i]]) *
           ((Transpose[invIi[a2, a3]].Ni[a2, a3])[[a, j]])) + c3[a2, a3] *
         (((invIi[a2, a3].Ni[a2, a3])[[i, a]]) * (IdentityMatrix[3][[j, b]]) -
           ((invIi[a2, a3].Ni[a2, a3].Ii[a2, a3])[[b, j]]) * (invIi[a2, a3][[i, a]]) +
           ((Transpose[invIi[a2, a3]].Ni[a2, a3])[[i, b]]) *
            (IdentityMatrix[3][[j, a]]) - ((Ii[a2, a3].Transpose[Ni[a2, a3]].
                invIi[a2, a3])[[j, a]]) * (invIi[a2, a3][[b, i]])), {i, 1, 3},
      {j, 1, 3}, {a, 1, 3}, {b, 1, 3}], Table[{b + 3 * (a - 1) + 9, j + 3 * (i - 1)} →
       -.5 * (((Transpose[invIi[a2, a3]].Ni[a2, a3])[[i, b]]) *
           (invIi[a2, a3][[j, a]])) -
        .5 * (((invIi[a2, a3].Ni[a2, a3])[[j, b]]) * (invIi[a2, a3][[a, i]])) +
        c3[a2, a3] *
         (-((invIi[a2, a3].Ni[a2, a3].Ii[a2, a3])[[j, b]]) * (invIi[a2, a3][[a, i]]) -
           ((Ii[a2, a3].Transpose[Ni[a2, a3]].invIi[a2, a3])[[b, i]]) *
            (invIi[a2, a3][[j, a]]) + ((invIi[a2, a3].Ni[a2, a3])[[a, i]]) *
            (IdentityMatrix[3][[j, b]]) + ((Transpose[invIi[a2, a3]].Ni[a2, a3])[[
              a, j]]) * (IdentityMatrix[3][[i, b]])), {i, 1, 3}, {j, 1, 3},
      {a, 1, 3}, {b, 1, 3}], Table[{b + 3 * (a - 1) + 9, j + 3 * (i - 1) + 9} →
       (.5 * invIi[a2, a3][[i, a]] + .5 * invIi[a2, a3][[a, i]]) *
         (IdentityMatrix[3][[b, j]]) +
        c3[a2, a3] * ((invIi[a2, a3][[i, a]]) * (Ii[a2, a3][[b, j]]) +
           (invIi[a2, a3][[a, i]]) * (Ii[a2, a3][[j, b]]) -
          2 * (IdentityMatrix[3][[a, j]]) * (IdentityMatrix[3][[b, i]])),
      {a, 1, 3}, {b, 1, 3}, {i, 1, 3}, {j, 1, 3}]]];
```

Okay, let's compute!

```
AbsoluteTiming[ltest = L[FFine[.56], .56]]
```

```
{94.290260,
 {{-0.240517, 0., 0., 0., 7.70355, 0., 0., 0., -4.25018, 0., -2.47912, 0., -38.4208,
   0., 0., 0., 0., 0.}, {0., 35.7632, 0., -5.99969, 0., 0., 0., 0., 0., 4.69132,
   0., 0., 0., -26.109, 0., 0., 0., 0.}, {0., 0., 14.8417, 0., 0., 0.,
   1.01436, 0., 0., 0., 0., 0., 0., 0., -19.2104, 0., -3.95267, 0.},
  {0., -5.99969, 0., 35.7632, 0., 0., 0., 0., 0., 4.69132, 0., 0., 0., -26.109,
   0., 0., 0., 0.}, {7.70355, 0., 0., 0., -9.15921, 0., 0., 0., -15.0688,
   0., -16.2862, 0., 142.857, 0., 0., 0., 0., 0.}, {0., 0., 0., 0., 0.,
   82.6, 0., -24.777, 0., 0., 0., -8.14308, 0., 0., 0., 40.926, 0., 0.},
  {0., 0., 1.01436, 0., 0., 0., 14.8417, 0., 0., 0., 0., 0., 0., 0., -19.2104,
   0., -3.95267, 0.}, {0., 0., 0., 0., 0., -24.777, 0., 82.6, 0., 0.,
   0., -8.14308, 0., 0., 0., 40.926, 0., 0.}, {-4.25018, 0., 0., 0.,
   -15.0688, 0., 0., 0., 16.2709, 0., 0., 0., 0., 0., 0., 0., 0., 0.},
  {0., 4.69132, 0., 4.69132, 0., 0., 0., 0., 0., 21.0635, 0., 0., 0., 0., 0., 0., 0., 0.},
  {-2.47912, 0., 0., 0., -16.2862, 0., 0., 0., 0., 0., 9.66041,
   0., 63.4625, 0., 0., 0., 0., 0.}, {0., 0., 0., 0., 0., -8.14308,
   0., -8.14308, 0., 0., 0., 1.16162, 0., 0., 0., 63.4625, 0., 0.},
  {-38.4208, 0., 0., 0., 142.857, 0., 0., 0., 0., 0., 63.4625, 0., -235.968,
   0., 0., 0., 0., 0.}, {0., -26.109, 0., -26.109, 0., 0., 0., 0., 0.,
   0., 0., 0., 0., 117.227, 0., 0., 0., 0.}, {0., 0., -19.2104, 0., 0.,
   0., -19.2104, 0., 0., 0., 0., 0., 0., 0., 6.46487, 0., 63.4625, 0.},
  {0., 0., 0., 0., 0., 40.926, 0., 40.926, 0., 0., 0., 63.4625, 0.,
   0., 0., -135.201, 0., 0.}, {0., 0., -3.95267, 0., 0., 0.,
   -3.95267, 0., 0., 0., 0., 0., 0., 0., 63.4625, 0., 30.8049, 0.},
  {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 67.1667}}}
```

Ouch, 90 seconds is a lot... Let's see if it is symmetric:

```
ltest - Transpose[ltest] // MatrixForm
```

```
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
```

Computation

```
AbsoluteTiming[JMar8 = Table[J[FFine[a3], a3], {a3, 0.0001, 1., .01}];]
```

```
{2171.142850, Null}
```

It took 31 minutes to compute J... Weird, for v4 it was 70... Something wrong?

```
xslot[slot_] := Take[Table[a3, {a3, 0.0001, 1., .01}], {slot}]
```

```
xslot[5]
```

```
{0.0401}
```

```
JMar8
```

```
JFeb25
```

```
JFeb3
```

```
Dimensions[JFeb3]
```

```
{100, 18, 18}
```

```
Jaugust3
```

```
Dimensions[Jaugust3]
```

```
{100, 18, 18}
```

```
Jaugust3list[slot_] := Flatten[Take[Jaugust3, {slot}], 1]
```

```
Jaugust3list[40] + Transpose[Jaugust3list[40]]
```

```
JFeb3list[slot_] := Flatten[Take[JFeb3, {slot}], 1]
```

```
JFeb3list[40] + Transpose[JFeb3list[40]]
```

```
JFeb25list[slot_] := Flatten[Take[JFeb25, {slot}], 1]
```

```
JMar8list[slot_] := Flatten[Take[JMar8, {slot}], 1]
```

Let's compare JFeb25 to JMar8, to see if there are any major differences:

```
JMar8list[40] - JFeb25list[40]
```

```
{{0., 0., 0., 0., 0., 0., 0., 0., 0.,
  -0.338086, 0., 0., 0., 0.0177454, 0., 0., 0., 0.0177454},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., -0.177916, 0., -0.0142552, 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., -0.177916, 0., 0., 0., -0.0270748, 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., -0.177916, 0., -0.0142552, 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0.0177454, 0., 0., 0.,
  -0.010765, 0., 0., 0., 0.0177454}, {0., 0., 0., 0., 0., 0., 0.,
  0., 0., 0., 0., 0., 0., 0., -0.0142552, 0., -0.0270748, 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., -0.177916, 0., 0., 0., -0.0270748, 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
  -0.0142552, 0., -0.0270748, 0.}, {0., 0., 0., 0., 0., 0., 0., 0.,
  0., 0.0177454, 0., 0., 0., 0.0177454, 0., 0., 0., -0.0364043},
 {0.338086, 0., 0., 0., -0.0177454, 0., 0., 0., -0.0177454, 0., 0.104555,
  0., -0.0687779, 0., 0., 0., 0., 0.}, {0., 0.177916, 0., 0.177916, 0.,
  0., 0., 0., 0., -0.104555, 0., 0., 0., 0.165632, 0., 0., 0., 0.0305382},
 {0., 0., 0.177916, 0., 0., 0., 0.177916, 0., 0., 0., 0., 0., 0., 0., 0.,
  0., 0.135093, 0.}, {0., 0.0142552, 0., 0.0142552, 0., 0., 0., 0.,
  0., 0.0687779, 0., 0., 0., -0.00770139, 0., 0., 0., 0.0305382},
 {-0.0177454, 0., 0., 0., 0.010765, 0., 0., 0., -0.0177454, 0., -0.165632,
  0., 0.00770139, 0., 0., 0., 0., 0.}, {0., 0., 0., 0., 0., 0.0142552, 0.,
  0.0142552, 0., 0., 0., 0., 0., 0., 0., 0., 0.0382396, 0., 0.}, {0., 0., 0.0270748,
  0., 0., 0., 0.0270748, 0., 0., 0., 0., 0., 0., 0., -0.0382396, 0., 0., 0.},
 {0., 0., 0., 0., 0., 0.0270748, 0., 0.0270748, 0., 0., 0., -0.135093,
  0., 0., 0., 0., 0., 0.}, {-0.0177454, 0., 0., 0., -0.0177454, 0., 0.,
  0., 0.0364043, 0., -0.0305382, 0., -0.0305382, 0., 0., 0., 0., 0.}}
```

There are! This should be interesting..

Okay, so, on the surface, J checks out. Now let's move to see if L checks out:

```
AbsoluteTiming[
 LMar8 = Table[L[FFine[a3], a3], {a3, 0.0001, 1., .01}];]
```

```
{7810.272054, Null}
```

It took 2.3 hours (8,351 seconds) to calculate L.

```
Laugust3
```

```
Dimensions[Laugust3]
```

```
{100, 18, 18}
```

```
LFeb3
```

```
LFeb25
```

```
LMar8
```

```
Laugust3list[slot_] := Flatten[Take[Laugust3, {slot}], 1]

LFeb3list[slot_] := Flatten[Take[LFeb3, {slot}], 1]

LFeb25list[slot_] := Flatten[Take[LFeb25, {slot}], 1]

LMar8list[slot_] := Flatten[Take[LMar8, {slot}], 1]

LFeb3list[40] - LFeb25list[40]

{{0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.},
 {0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.}}
```

Okay! So there are a few more discrepancies but still low orders of magnitude...

Eigenvalues of J.L

Now let's combine the two:

```
JLMar8[slot_] := JMar8list[slot].LMar8list[slot]

Dimensions[JLFeb25[4]]

{18, 18}

a22[slot_] := FFine[0.01 * (slot - 1) + 0.0001]
a33[slot_] := 0.01 * (slot - 1) + 0.0001
```

Now, recall that the equation is of the form:

$$\dot{\delta z} = J.L\,\delta z = \lambda \delta z,$$

so the solutions are of the form:

$\delta z = e^{\lambda t}$.

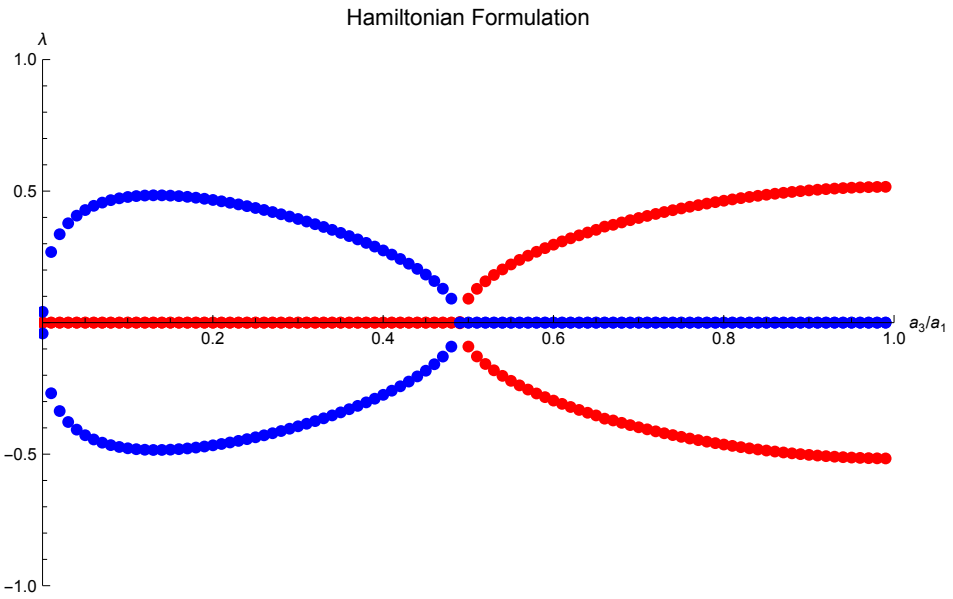
Hamiltonian systems always have "conjugate" modes - that is, for every eigenvalue, there is another one that is its negative (or conjugate if purely imaginary). This is a consequence of Liouville's theorem (if not, there would could be two unstable directions and phase space volume would increase). Therefore, any real eigenvalue is going to have a positive and negative value. **Therefore, we are looking for purely imaginary.**

```
Transpose[Table[Chop[Eigenvalues[JLMar8[x]], 10^-10], {x, 5, 100, 1}]] // MatrixForm
```

0. + 2.04649 i	0. + 2.05299 i	0. + 2.05872 i	0. + 2.06385 i	0. + 2.06849 i	0. + 2.
0. − 2.04649 i	0. − 2.05299 i	0. − 2.05872 i	0. − 2.06385 i	0. − 2.06849 i	0. − 2.
0. + 1.63588 i	0. + 1.64499 i	0. + 1.65386 i	0. + 1.66261 i	0. + 1.67132 i	0. + 1.
0. − 1.63588 i	0. − 1.64499 i	0. − 1.65386 i	0. − 1.66261 i	0. − 1.67132 i	0. − 1.
0. + 1.45734 i	0. + 1.45849 i	0. + 1.45826 i	0. + 1.45701 i	0. + 1.45501 i	0. + 1.
0. − 1.45734 i	0. − 1.45849 i	0. − 1.45826 i	0. − 1.45701 i	0. − 1.45501 i	0. − 1.
0. + 1.34988 i	0. + 1.34095 i	0. + 1.33302 i	0. + 1.32585 i	0. + 1.31929 i	0. + 1.
0. − 1.34988 i	0. − 1.34095 i	0. − 1.33302 i	0. − 1.32585 i	0. − 1.31929 i	0. − 1.
0.406698	0.428007	−0.444086	−0.456337	−0.465647	0.47
−0.406698	−0.428007	0.444086	0.456337	0.465647	−0.4
0. + 0.0609865 i	0. + 0.0763467 i	0. + 0.0913314 i	0. + 0.105918 i	0. + 0.120093 i	0. + 0.
0. − 0.0609865 i	0. − 0.0763467 i	0. − 0.0913314 i	0. − 0.105918 i	0. − 0.120093 i	0. − 0.
0	0	0	0	0	
0	0	0	0	0	
0	0	0	0	0	
0	0	0	0	0	
0	0	0	0	0	
0	0	0	0	0	

```
imeigen[y_] := Table[Flatten[
    {xslot[x], Chop[Im[Eigenvalues[JLMar8[x]]], 10^-10][[y]]}, 1], {x, 1, 100}];

reeigen[y_] := Table[Flatten[
    {xslot[x], Chop[Re[Eigenvalues[JLMar8[x]]], 10^-10][[y]]}, 1], {x, 1, 100}];

imeigen1[y_] := Table[Flatten[
    {xslot[x], Chop[Im[Eigenvalues[JLMar8[x]]], 10^-10][[y]]}, 1], {x, 1, 35}];

reeigen1[y_] := Table[Flatten[
    {xslot[x], Chop[Re[Eigenvalues[JLMar8[x]]], 10^-10][[y]]}, 1], {x, 1, 35}];

imeigen2[y_] := Table[Flatten[
    {xslot[x], Chop[Im[Eigenvalues[JLMar8[x]]], 10^-10][[y]]}, 1], {x, 36, 66}];

reeigen2[y_] := Table[Flatten[
    {xslot[x], Chop[Re[Eigenvalues[JLMar8[x]]], 10^-10][[y]]}, 1], {x, 36, 66}];

imeigen3[y_] := Table[Flatten[
    {xslot[x], Chop[Im[Eigenvalues[JLMar8[x]]], 10^-10][[y]]}, 1], {x, 67, 100}];

reeigen3[y_] := Table[Flatten[
    {xslot[x], Chop[Re[Eigenvalues[JLMar8[x]]], 10^-10][[y]]}, 1], {x, 67, 100}];
```

```
Show[ListPlot[{imeigen1[10], imeigen1[9], reeigen1[10], reeigen1[9]},
  PlotStyle → {Red, Red, Blue, Blue}, PlotRange → {{0, 1}, {-1, 1}},
  AxesLabel → {"a₃/a₁", "λ"}, PlotLabel → "Hamiltonian Formulation"],
 ListPlot[{imeigen2[11], imeigen2[12], reeigen2[11], reeigen2[12]},
  PlotStyle → {Red, Red, Blue, Blue}, PlotRange → {-1, 1}, AxesLabel → {"a₃/a₁", "λ"}],
 ListPlot[{imeigen3[10], imeigen3[9], reeigen3[10], reeigen3[9]},
  PlotStyle → {Red, Red, Blue, Blue}, PlotRange → {-1, 1}, AxesLabel → {"a₃/a₁", "λ"}]]
```

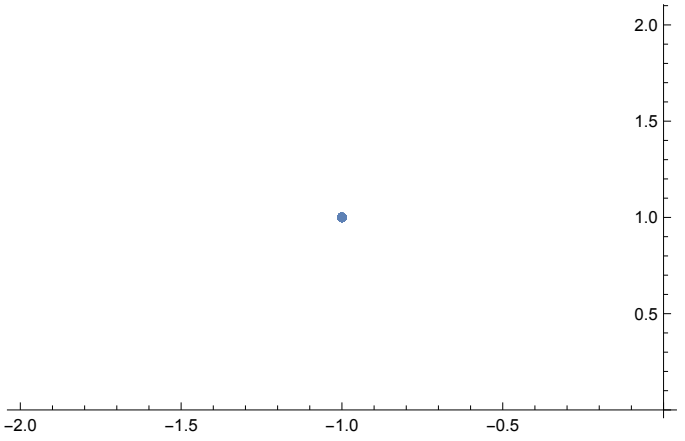


So the units didn't make a difference... That is, they weren't introducing any error since we are getting stability at the same $\tilde{a}_3$ as in the previous version of this *Mathematica* notebook. It seems like this is more systematic :(. Where is the problem!?

Definitiveness of L

The definitiveness of L is determined by seeing if all the eigenvalues are of the same sign. Let's see if we can check this somehow.

```
ListPlot[Table[{Min[Sign[Eigenvalues[LMar8list[x]]]],
   Max[Sign[Eigenvalues[LMar8list[x]]]]}, {x, 1, 100}]]
```



And I'm getting that L is totally indefinite...

```
Transpose[Table[Chop[Eigenvalues[LFeb25list[x]], 10^-10], {x, 5, 100, 10}]] //
 MatrixForm
```

2.53504×10^6	24 017.9	3006.21	732.556	242.725	−114.426	−92.2823	−99.8698	−16
1.84695×10^6	14 176.5	1661.21	402.65	−180.153	95.5825	−50.8873	−71.9715	−14
789 629.	−5119.45	−1003.86	−361.204	138.895	58.9823	41.8289	−59.3345	−13
−229 418.	4780.49	550.563	156.537	78.2649	45.327	−29.2502	19.5357	11.
14 749.2	1100.37	336.661	148.958	66.6769	−44.1307	28.3237	17.8033	9.4
10 405.4	603.536	183.922	86.5783	52.9096	34.1163	27.8757	14.6113	9.4
4342.48	427.336	135.871	83.919	47.6987	33.5247	22.285	14.5934	9.0
3109.43	254.738	127.407	−47.363	−43.5509	28.7485	18.5757	12.6944	8.4
−3100.94	190.489	86.7333	44.9216	25.8147	17.1404	18.246	11.3287	8.2
−3012.33	113.994	−57.5929	43.2272	16.6874	11.0468	12.2032	9.12829	7.0
876.505	98.8285	50.2986	27.9569	15.8607	10.9046	8.10968	6.6132	5.8
748.78	88.2771	46.6231	25.4162	15.368	10.8248	8.05532	6.06571	4.6
330.713	−85.9155	35.3054	21.6325	−13.2137	−9.58335	7.50136	5.38925	4.5
−263.446	−35.8124	28.2316	−19.7088	12.949	−8.8156	5.97359	4.74535	3.2
−18.9566	5.3361	−28.0515	19.2563	10.3792	8.37827	−5.91312	4.38775	−2.
6.69914	4.89951	5.54515	5.71188	5.81646	5.8317	5.7167	−3.95825	2.5
5.09961	−4.79233	4.78333	4.63217	4.4243	4.11243	3.62058	2.92613	2.1
4.98386	4.62887	3.87487	3.1619	2.42252	1.68041	1.02078	0.51223	0.1

Comparison to Chandrasekhar's 1965 paper

Characteristic equation is:

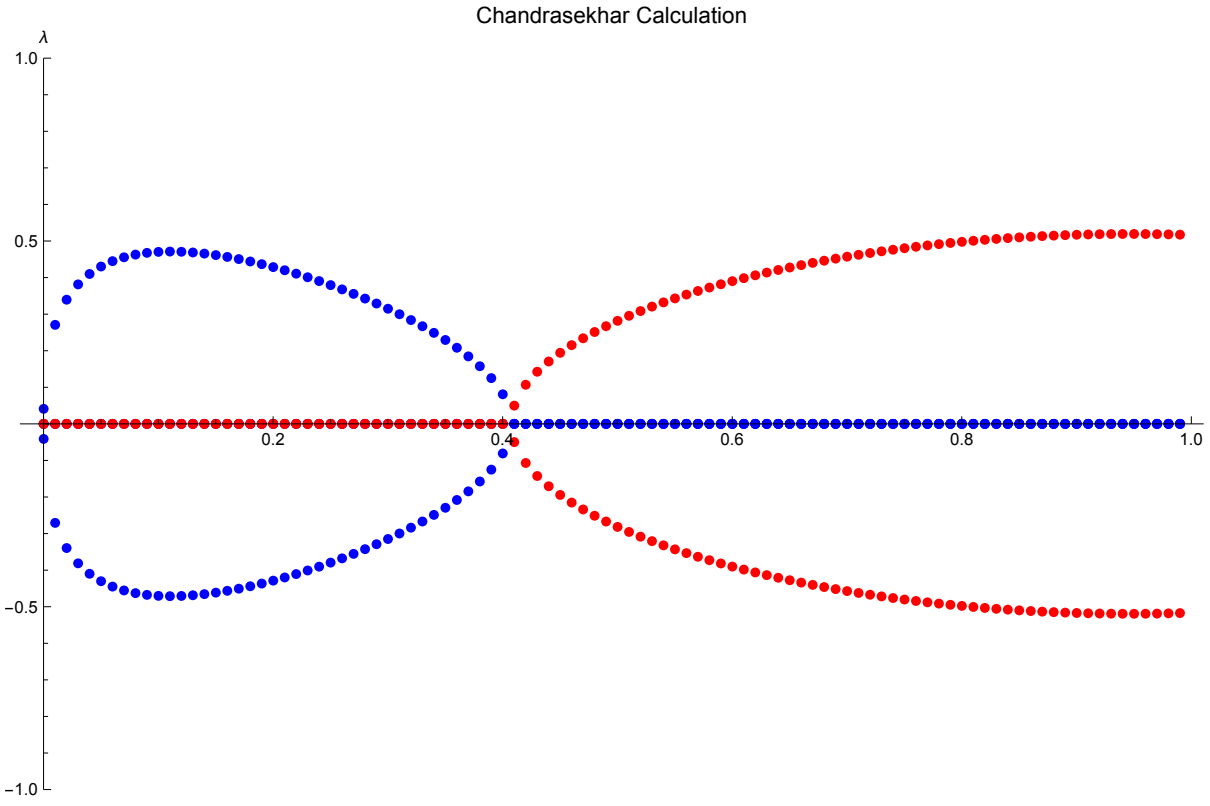
$$\lambda^4 + (4\,B_{13} + 4\,B_{23} + 2\,B_{12})\,\lambda^2 + (4\,B_{13} - \Omega\,(K_{12} + \Omega))\,(4\,B_{23} + \Omega(K_{21} - \Omega)) = 0$$

Also note that $Q_1 = (K_{12} + \Omega)$ and $Q_2 = (K_{21} - \Omega)$.

So:

```
B13[a2_, a3_] :=
  NIntegrate[a2 * a3 * u / (Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (1 + u) * (a3^2 + u)), {u, 0, Infinity}];
B23[a2_, a3_] :=
  NIntegrate[a2 * a3 * u / (Sqrt[(1 + u) (a2^2 + u) (a3^2 + u)] * (a3^3 + u) * (a2^2 + u)), {u, 0, Infinity}];
a3 =.
```

```
realeigenvalsaug12 =
  Table[Partition[Riffle[Table[a3, {a3, 0.0001, 1., .01}], Table[Re[x /. Solve[
          x^4 + (4 B13[FFine[a3], a3] + 4 B23[FFine[a3], a3] + 2 B12[FFine[a3], a3])
             x^2 + (4 * B13[FFine[a3], a3] - (Ω3[FFine[a3], a3]) * (Kt[FFine[a3], a3][[
                    1, 2]] + Ω3[FFine[a3], a3])) * (4 * B23[FFine[a3], a3] + (Ω3[FFine[
                    a3], a3]) * (Kt[FFine[a3], a3][[2, 1]] - Ω3[FFine[a3], a3])) ==
           0, x]][[i]], {a3, 0.0001, 1., .01}]], 2], {i, 1, 4}];
imeigenvalsaug12 = Table[Partition[Riffle[Table[a3, {a3, 0.0001, 1., .01}],
     Table[Im[x /. Solve[
          x^4 + (4 B13[FFine[a3], a3] + 4 B23[FFine[a3], a3] + 2 B12[FFine[a3], a3])
             x^2 + (4 * B13[FFine[a3], a3] - (Ω3[FFine[a3], a3]) * (Kt[FFine[a3], a3][[
                    1, 2]] + Ω3[FFine[a3], a3])) * (4 * B23[FFine[a3], a3] + (Ω3[FFine[
                    a3], a3]) * (Kt[FFine[a3], a3][[2, 1]] - Ω3[FFine[a3], a3])) ==
           0, x]][[i]], {a3, 0.0001, 1., .01}]], 2], {i, 1, 4}];

ListPlot[{Flatten[realeigenvalsaug12, 1], Flatten[imeigenvalsaug12, 1]},
 AxesLabel → {"ã₃", "λ"}, PlotRange → {-1, 1},
 PlotStyle → {Blue, Red}, PlotLabel → "Chandrasekhar Calculation"]
```




```
Export["chandraeigen.eps", Out[74]]
```

chandraeigen.eps

Bibliography

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

[2] Audin, M. (1996). *Spinning Tops: A Course on Integrable Systems*. Cambridge University Press.

[3] Chandrasekhar, S. (1965). The equilibrium and stability of the riemann ellipsoids. *Ap. J.*, 142:890–921.

[4] Chandrasekhar, S. (1969). *Ellipsoidal Figures of Equilibrium*. Yale University Press, New Haven, 1 edition.

[5] Chandrasekhar, S. and Lebovitz, N. R. (1962). On the oscillations and the stability of rotating gaseous masses. *Ap. J.*, 135:248–260.

[6] Chandre, C., de Guillebon, L., Back, A., Tassi, E., and Morrison, P. J. (2013). On the use of projectors for hamiltonian systems and their relationship with dirac brackets. *J. Phys. A: Math. Theor.*, 46.

[7] Chandre, C., Morrison, P. J., and Tassi, E. (2012). On the hamiltonian formulation of incompressible ideal fluids and magnetohydrodynamics via dirac's theory of constraints. *Phys. Lett. A*, 376:737–743.

[8] Cohen, H. and Muncaster, R. G. (1988). The theory of pseudo-rigid bodies. *Springer Tracts in Natural Philosophy*, 33.

[9] da Silva, A. C. (2011). *Introduction to Symplectic and Hamiltonian Geometry*.

[10] Dyson, F. J. (1968). Dynamics of a spinning gas cloud. *J. Math. Mech.*, 18.

[11] Fassò, F. and Lewis, D. (2001). Stability properties of the riemann ellipsoids. *Arch. Rational Mech. Anal.*, 158:259–292.

[12] Guillemin, V. and Pollack, A. (1974). *Differential Topology*. Englewood Cliffs, N.J.: Prentice-Hall.

[13] Hirota, M., Morrison, P. J., and Hattori, Y. (2014). Variational necessary and sufficient stability conditions for inviscid shear flow. *Proc. R. Soc. Lond. Ser. A*, 470.

[14] Kundu, P. and Cohen, I. M. (2002). *Fluid Mechanics*. Elsevier Science (USA), 2 edition.

[15] Landau, L. D. and Lifshitz, E. M. (1987). *Fluid Mechanics*. Pergamon Press, 2 edition.

[16] Lebovitz, N. R. (1961). The virial tensor and its application to self-gravitating fluids. *Ap. J.*, 134:500–536.

[17] Lebovitz, N. R. (1966). On riemann's criterion for the stability of liquid ellipsoids.'. *Ap. J.*, 145:878–885.

[18] Lebovitz, N. R. (1972). On the fission theory of binary stars. *Ap. J.*, 175:171–183.

[19] Lebovitz, N. R. and Lifschitz, A. (1996a). New global instabilities of the riemann ellipsoids. *Ap. J.*, 458:699–713.

[20] Lebovitz, N. R. and Lifschitz, A. (1996b). Short-wavelength instabilities of riemann ellipsoids. *Phil. Trans. Roy. Soc. London A*, 354:927–950.

[21] Lee, J. M. (2013). *Introduction to Smooth Manifolds*. Springer-Verlag New York Inc., 2 edition.

[22] Lewis, D. and Simo, J. C. (1990). Nonlinear stability of rotating pseudo-rigid bodies. *Proc. R. Soc. Lond. Ser. A*, 427:281–319.

[23] Lyttleton, R. A. (1953). *The Stability of Rotating Liquid Masses*. Cambridge University Press.

[24] Morrison, P. J. (1998). Hamiltonian description of the ideal fluid. *Rev. Mod. Phys.*, 70:468–519.

[25] Morrison, P. J. and Greene, J. M. (1980). Noncanonical hamiltonian density formulation of hydrodynamics and ideal magnetohydrodynamics. *Phys. Rev. Lett.*, 45:790–793.

[26] Morrison, P. J. and Hagstrom, G. I. (2014). *Nonlinear Physical Systems: Spectral Analysis, Stability and Bifurcations (Chapter 12)*. John Wiley & Sons, Inc.

[27] Morrison, P. J., Lebovitz, N. R., and Biello, J. A. (2009). The hamiltonian description of incompressible fluid ellipsoids. *Ann. Phys.*, 324:1747–1762.

[28] Riemann, B. (1861). Ein beitrag zu den untersuchungen über die bewegung cines flüssigen gleichartigen ellipsoides. *Abh. d. Königl. Gesell. der Wis. zu Göttingen*, 9:168–197.

[29] Rosensteel, G. (1988). Rapidly rotating nuclei as riemann ellipsoids. *Ann. Phys.*, 186:230–291.

[30] Salmon, R. (1988). Hamiltonian fluid mechanics. *Ann. Rev. Fluid Mech.*, 20:225–256.

[31] Sundermeyer, K. (1982). *Constrained Dynamics*. Springer-Verlag New York Inc.

[32] Tritton, D. J. (2007). *Physical Fluid Dynamics*. Oxford University Press, 2 edition.

[33] Whymark, A. Primary shape formation: The splashform morphology. http://www.tektites.co.uk/primary-shapes.html.