Statistical mechanics of self-gravitating systems

Jun Yan Lau

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Space and Climate Physics

University College London

September 8, 2024

I, Jun Yan Lau, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

It is widely known that the notion of maximum entropy states describing macroscopic steady-states for many-body systems does not apply for astrophysical systems. A theory of van Kampen modes contextualised to stellar systems is presented which captures fluctuations as solutions to the linearised Collisionless Boltzmann Equation (CBE). The CBE is elevated from an algebraic equation with derivatives in phase-space to a functional equation acting with functional derivatives, and finally a statistical mechanical theory defining ensemble averages in a physics-agnostic manner is presented and applied to self-gravitating systems. Correlation functions are computed from this theory that explain how gravitational dressing causes self-gravitating systems to depart from the maximum entropy state, and will be crucial to fully resolve the substructure found in the highly precise Gaia photometric data.

Impact Statement

This thesis challenges the conventional usage of the principle of maximum entropy as a way of ascertaining the state of maximum likelihood by pointing out that the principle of maximum entropy only constrains the probabilistic distribution of particles, and not the true distribution of particles that can be obtained by sampling the probability distribution.

This result has widespread applications in the field of information theory, where I show that maximising entropy assumes that all the information inherent to the system is stored in the distribution function, and not the actual particles that form the true system. This assumption fails when the particles are correlated in any non-trivial manner. My work establishes how we can maximise a different entropy—one that combines the information content of the distribution and the information content of the sample obtained from the distribution into a single monolithic whole, so as to include these strong correlations.

UCL Research Paper Declaration Form: referencing the doctoral candidate's own published work(s): Paper 1

1. **1. For a research manuscript that has already been published** (if not yet published, please skip to section 2)**:**

   (a) **What is the title of the manuscript?** Modes of a Stellar System I: ergodic systems

   (b) **Please include a link to or doi for the work:** DOI: 10.1093/mnras/stab2227

   (c) **Where was the work published?** Monthly Notices of the Royal Astronomical Society

   (d) **Who published the work?** Oxford Univeristy Press

   (e) **When was the work published?** 02 August 2021

   (f) **List the manuscript's authors in the order they appear on the publication:** Jun Yan Lau, James Binney.

   (g) **Was the work peer reviewd?** Yes

   (h) **Have you retained the copyright?** Yes

   (i) **Was an earlier form of the manuscript uploaded to a preprint server (e.g. medRxiv)? If 'Yes', please give a link or doi** Yes, DOI: 10.48550/arXiv.2104.07044

   If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

   ☐ *I acknowledge permission of the publisher named under 1d to include in this thesis portions of the publication named as included in 1c.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)**:**

(a) **What is the current title of the manuscript?**

(b) **Has the manuscript been uploaded to a preprint server 'e.g. medRxiv'?**

   If 'Yes', please please give a link or doi:

(c) **Where is the work intended to be published?**

(d) **List the manuscript's authors in the intended authorship order:**

(e) **Stage of publication:**

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)**:** Jun Yan Lau did the theorycrafting and described part of the interpretation of the mathematics into physics, and put it in writing. James Binney refined and extended the interpretation and writing.

4. **In which chapter(s) of your thesis can this material be found?** Chapter 2.

   e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)**:**

Candidate: Jun Yan Lau

Date:01/07/2024

Supervisor/Senior Author signature (where appropriate)**:**Ralph Schönrich

Date:01/07/2024

UCL Research Paper Declaration Form: referencing the doctoral candidate's own published work(s): Paper II

1. **1. For a research manuscript that has already been published** (if not yet published, please skip to section 2)**:**

   (a) **What is the title of the manuscript?** Modes of a stellar system II: non-ergodic systems

   (b) **Please include a link to or doi for the work:** DOI: 10.1093/mnras/stab2175

   (c) **Where was the work published?** Monthly Notices of the Royal Astronomical Society

   (d) **Who published the work?** Oxford University Press

   (e) **When was the work published?** 29 July 2021

   (f) **List the manuscript's authors in the order they appear on the publication:** Jun Yan Lau, James Binney

   (g) **Was the work peer reviewed?** Yes

   (h) **Have you retained the copyright?** Yes

   (i) **Was an earlier form of the manuscript uploaded to a preprint server (e.g. medRxiv)? If 'Yes', please give a link or doi** DOI: 10.48550/arXiv.2106.10297

      If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

      ☐ *I acknowledge permission of the publisher named under 1d to include in this thesis portions of the publication named as included in 1c.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)**:**

(a) **What is the current title of the manuscript?**

(b) **Has the manuscript been uploaded to a preprint server 'e.g. medRxiv'?**

   If 'Yes', please please give a link or doi:

(c) **Where is the work intended to be published?**

(d) **List the manuscript's authors in the intended authorship order:**

(e) **Stage of publication:**

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)**:** Jun Yan Lau did the theorycrafting and described part of the interpretation of the mathematics into physics, and put it in writing. James Binney refined and extended the interpretation and writing.

4. **In which chapter(s) of your thesis can this material be found?** Chapter 3

   e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)**:**

Candidate: Jun Yan Lau

Date:01/07/2024

Supervisor/Senior Author signature (where appropriate)**:**Ralph Schönrich

Date: 01/07/2024

UCL Research Paper Declaration Form: referencing the doctoral candidate's own published work(s): Paper III

1. **1. For a research manuscript that has already been published** (if not yet published, please skip to section 2)**:**

   (a) **What is the title of the manuscript?** Probabilistic distribution functions

   (b) **Please include a link to or doi for the work:** DOI: 10.1093/mnras/stab2047

   (c) **Where was the work published?** Monthly Notices of the Royal Astronomical Society

   (d) **Who published the work?** Oxford University Press

   (e) **When was the work published?** 16 July 2021

   (f) **List the manuscript's authors in the order they appear on the publication:** Jun Yan Lau, James Binney

   (g) **Was the work peer reviewd?** Yes

   (h) **Have you retained the copyright?** Yes

   (i) **Was an earlier form of the manuscript uploaded to a preprint server (e.g. medRxiv)? If 'Yes', please give a link or doi** DOI: 10.48550/arXiv.2107.06298

   If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

   ☐ *I acknowledge permission of the publisher named under 1d to include in this thesis portions of the publication named as included in 1c.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)**:**

(a) **What is the current title of the manuscript?**

(b) **Has the manuscript been uploaded to a preprint server 'e.g. medRxiv'?**

   If 'Yes', please please give a link or doi:

(c) **Where is the work intended to be published?**

(d) **List the manuscript's authors in the intended authorship order:**

(e) **Stage of publication:**

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)**:** Jun Yan Lau did the theorycrafting and described part of the interpretation of the mathematics into physics, and put it in writing. James Binney refined and extended the interpretation and writing.

4. **In which chapter(s) of your thesis can this material be found?** Chapter 4

   e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)**:**

Candidate: Jun Yan Lau
Date:01/07/2024

Supervisor/Senior Author signature (where appropriate)**:**Ralph Schönrich
Date:01/07/2024

UCL Research Paper Declaration Form: referencing the doctoral candidate's own published work(s): Paper IV

1. **1. For a research manuscript that has already been published** (if not yet published, please skip to section 2)**:**

   (a) **What is the title of the manuscript?**

   (b) **Please include a link to or doi for the work:**

   (c) **Where was the work published?**

   (d) **Who published the work?**

   (e) **When was the work published?**

   (f) **List the manuscript's authors in the order they appear on the publication:**

   (g) **Was the work peer reviewed?**

   (h) **Have you retained the copyright?**

   (i) **Was an earlier form of the manuscript uploaded to a preprint server (e.g. medRxiv)? If 'Yes', please give a link or doi**

   If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

   ☐ *I acknowledge permission of the publisher named under 1d to include in this thesis portions of the publication named as included in 1c.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)**:**

   (a) **What is the current title of the manuscript?** Boltzmann Equation Field Theory I: Ensemble Averages

(b) **Has the manuscript been uploaded to a preprint server 'e.g. medRxiv'?**

   If 'Yes', please please give a link or doi: Yes, DOI: 10.48550/arXiv.2310.04508

(c) **Where is the work intended to be published?** Monthly Notices of the Royal Astronomical Society

(d) **List the manuscript's authors in the intended authorship order:** Jun Yan Lau

(e) **Stage of publication:** Submitted

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)**:**

4. **In which chapter(s) of your thesis can this material be found?** Chapter 5

   e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)**:**

Candidate: Jun Yan Lau

Date: 01/07/2024

Supervisor/Senior Author signature (where appropriate)**:**

Date:

Acknowledgements

Thanks to my parents, Lau Weng Onn and Low Whai Mun, without whom I would never have embarked on my research career: your encouragement and unending supply of (read: sometimes irritating) wisdom which stood strong even when I didn't see the big picture means the world to me. They saw who I could be before I did.

Thanks to my 'sciencey seniors', my supervisors, tutors and friends, Ralph Schoenrich, James Binney, John Magorrian, Alexander Schekochihin, Justin Wark, Peter Read, Francesco Hautmann and last but not least, Sam Vinko, without whom I would never have understood what a research career should look like, how a scientist should conduct themselves (and I would probably be out of a job in science too!): you set a high bar worthy of striving to reach, and helped me temper and qualify my desire for success. And of course, thanks for the enjoyable banter. After all, what's the point of having the gift of the gab if you don't have someone to gabber along with?

Thanks to my friends, who listened to painful hours of my droning on and on about my work. I enjoy what I do, and a great deal of that is thanks to you. Watching you writhe in pain at arcane words like "astrophysics" and "maximum entropy" was fun, but it is those moments in which we genuinely shared ideas not just about work but also about life, that I am truly grateful for.

Thanks to my girlfriend, Liu Jingting for inspiring me to put my heart into everything I do. This thesis would be much more of a boring read without you—after all, it's hard to get excited about writing a 150 page long formality.

How about a 150 page long labour of love? That I can do.

Contents

Chapter 1

Introduction

The goal of statistical mechanics is to connect the macroscopic features of systems with the microscopic interactions between the particles that comprise them. In the context of astrophysical dynamics, the macroscopic features are the morphologies and dynamics of globular clusters and galaxies, while the microscopic interactions are the gravitational forces between the stars that comprise these systems.

This thesis describes my efforts in creating a theory that attempts to bridge an important gap in the field of astrophysical dynamics: What is the relationship between real astrophysical systems that are composed of N discrete stars interacting via the gravitational force law to the approximations dynamicists usually use to model them: the phase-fluid that flows under its own gravity?

Or more to the point, how do we connect the macroscopic and microscopic features of self-gravitating systems? The statistical mechanics of self-gravitating systems will be an important tool to describe the out-of-equilibrium, unsmooth nature of modern observations of astrophysical systems.

To begin with, let us ask the question: What are microscopic and macroscopic features?

1.1   The History of Statistical Mechanics

A microstate is a list of positions and velocities, phase-space coordinates $\mathbf{w} = (\mathbf{x}, \mathbf{v})$ of length N, $\{\mathbf{w}_i\} \equiv \{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_N\}$ describing a system at a single point in time.

A macrostate on the other hand, is defined in a self-referential manner. A macrostate is defined as a complete list of macroscopic variables—variables that are thermodynamically relevant to a system, while a theory of thermodynamics is defined by how it relates one macroscopic variable to another; as exemplified by the first law of thermodynamics (Clausius 1850); $\Delta U = Q - W$ that relates the change in the internal energy ΔU of a closed system with the difference of the heat Q introduced into the system and the thermodynamic work W done by the system. The reason macroscopic variables are defined in this cyclical manner is that thermodynamics was developed first as a phenomenological theory, before it was described by statistical mechanics. Defining the concept of a macroscopic feature thus requires the inspection of a thermodynamical theory.

Classical thermodynamics, which arises from the microscopic interpretation provided by Boltzmann-Gibbs statistical mechanics (BGSM), prescribes relationships between macroscopic quantities such as pressure, temperature and work done under certain assumptions.

These macroscopic quantities all share a common definition that was first captured by Daniel Bernoulli (in his Hydrodynamica, Chapter 10, Sections 4 and 6, where he outlines the kinetic theory of gases) who singled out a class of macroscopic quantities by focusing on the ones that can be understood by time-averaging their corresponding microscopic quantities—pressure from taking a time average of the momentum transferred from gas particles hitting the walls of a box, for example. A calculation like this, however, requires integrating an initial condition of microstates forwards in time, and was intractable for interacting systems.

It was Gibbs (1902) who explicitly replaced the deterministic but chaotic

production of microstates sourced from time-evolution with a stochastic alternative. He posits that on the lengthy time-scales over which Bernoulli's macroscopic quantities are measured, the particles have had sufficient opportunity to 'rearrange' themselves, their time-evolution sampling all the microstates available to them with equal probability: the so-called ergodic hypothesis.

The ergodic hypothesis is a justification and mathematical encoding of one of Boltzmann's postulates: that the dynamics of microstates functions to chaotically (in the classical sense of mixing) 'scramble' information that is inherent to the microstate. This mixing conserves only the collisional invariants while maximising uncertainty regarding our knowledge of the microstates; a postulate he used to derive the Maxwellian (thus, Maxwell-Boltzmann) distribution. This postulate (and thus the ergodic hypothesis) is well suited for a gas with particles that exhibit short-ranged interactions, where interactions deflect particles from their original trajectories thus 'scrambling' the system but conserving total momentum and energy. However, does it describe systems with strong, long-ranged interactions? Gibbs certainly did not think so, thus he provided an additional assumption: that the energy of the system $E[\{\mathbf{w}_i\}] = \sum_i E(\mathbf{w}_i)$ could be expressed as a sum of the energy of each particle within the microstate, or that systems described by BGSM had to be composed of weakly-interacting particles.

The ergodic hypothesis aligns with the intuition that one should only measure the pressure of a system by summing over a large number of collisions— so as to suppress noise fluctuations in the collision rate associated with the stochastic nature of particles. It not only does away with the need to tackle (weak) dynamics, but also removes the need for a microstate. Hence it is understood that BGSM only applies for systems that are at equilibrium/adiabatically changing; that is systems for which the macroscopic variables change far more slowly than the time taken for a particle to make its rounds within the system, which is a function of the speed of sound and the size of the system considered.

It is worth noting that while BGSM describes much of classical thermodynamics related to gases; it does not cover phase-transitions or criticality: neither does it cover solids or fluids. Those applications are also beyond the scope of this thesis, given that we are interested in a gas of stars.

What fundamentally prevents applications of BGSM to astrophysical systems is not just that it only applies to weakly interacting systems with short ranged forces (the gravitational force is long-ranged and therefore creates a strongly interacting system), or just that the ergodic hypothesis does not function (the surface of constant energy defined in the space of microstates is unbounded in self-gravitating systems, and hence the ergodic hypothesis fails), but it is that the macroscopic features we astrophysicists are interested in are not described by BGSM.

When an inherently chaotic (but weakly interacting) system evolves for a sufficiently long period, time averages naturally equate to ensemble averages, which causes measurements that are taken over such periods to correspond to ensemble averages (i.e. a measurement of temperature). However, does this picture align with the way we observe stars? Observations of stars within our Milky Way are made within the slightest of instants, relative to astrophysical timescales. We are in radically different regimes to Gibbs and Bernoulli: our macroscopic features are not ones which are persistent such that they evolve only across secular timescales, but rather are system-scale phase-space fluctuations—They are collective motions in the microstate: spirals, bars, and dipole asymmetries which do not belong under the category of Bernoulli's macroscopic quantities. This is the difference between a measurement of the velocity-dispersion and a measurement of the temperature of a system: measuring the latter implies stationarity; while you can measure the velocity-dispersion at any point in time.

To address the question that kickstarted this investigation into thermodynamics, I propose that a macroscopic feature is a common feature found amongst all representative models of a system—this captures both tempo-

rally persistent features and features with strong phase-space signatures and removes human bias in defining macroscopic quantities.

This novel definition of a macroscopic feature allows us to refine the definition of the macrostate. The macrostate is thus defined as a complete list of all common features found amongst all representative models of a system, and is therefore the distribution of all representative models of a system.

Now we have one question left to answer: How do we (representatively) model an N-particle self-gravitating system?

1.2   The Collisionless Boltzmann Equation

Owing to the belief that BGSM doesn't apply to astrophysical systems, the statistical mechanics approach has generally been eschewed when it comes to predicting the macroscopic features of self-gravitating systems in favour of the application of dynamics

So how do we understand (and therefore model) the evolution of a system comprising N stars that evolves under the gravitational force law from an initial microstate?

While we could integrate $6N$ coupled ordinary differential equations (the force equations) forwards in time, this is computationally expensive, becoming intractable for galaxies where the number of stars, $N \sim 10^{11}$, and that is before we include the dark matter contribution. We know each star gravitationally interacts every other star, but then what else can we glean from a list of phase-space coordinates 10^{11} long? To obtain an understanding of the dynamics of these systems, it is customary to view them via histograms or scatter diagrams: any method that emphasises the density of stars in a region, (e.g. Antoja et al. (2018)). We are not so much interested in the phase-space coordinates of each and every star in a galaxy as we are interested in the phase-space density of stars: we want to understand the dynamics of the bulk.

We can construct a statistically smooth distribution of stars from a microstate via smoothening: taking phase-space averages of the number of stars (the occupancy number, n) within regions ensuring that n is large enough to suppress Poisson noise statistics. One can also spread each star's mass out over a distance several times larger than the local interstellar distance in phase-space, achieving similar results.

We might also prefer not to create a smooth distribution: instead deciding that we want a coarse-grained DF, that is a DF defined by carving phase-space into individual elements and then setting the value of the DF to be the average value within each element. While it is easier to obtain from observations (just count particles in each element) and easier to interpret (there can be no

features finer than the interstellar separation) the difficulty of this option lies in how there are no obvious evolution equations for these systems. Tremaine et al. (1986) laid down the foundations by connecting these coarse-grained DFs with their smooth equivalents, while Dehnen (2005) extended their conception of 'mixing'. Perhaps more recently, Barbieri et al. (2022) described a way of coarse-graining while preserving symplectic structure—though only for strongly restricted systems, and producing mixed results in comparisons with their simulations.

Regardless of the manner in which it is obtained, the value the phase space density f takes in a region must approximate the average number of stars frequenting a unit phase-space volume in that region of phase-space. Thus the phase space density f is commonly interpreted as the probabilistic distribution function (DF) describing the density of stars in a region. This connection is formalised by Liouville's theorem.

Liouville's theorem describes the conservation of a phase-fluid DF along the trajectory of the system. For an N-particle system, we may define the N-particle phase-space DF $f^{(N)}(\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_N) = f^{(N)}[\{\mathbf{w}_i\}]$ that evolves like a Hamiltonian system. Then Liouville's theorem is:

$$\frac{\partial f^{(N)}}{\partial t} + \sum_i \dot{\mathbf{w}}_i \frac{\partial f^{(N)}}{\partial \mathbf{w}_i} = 0 \tag{1.1}$$

where $\dot{\mathbf{w}}_i$ is defined by Hamilton's equations, $\mathbf{w} = (\mathbf{q}, \mathbf{p})$ are canonical coordinates such that:

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\partial H}{\partial \mathbf{p}_i} \\ \dot{\mathbf{p}}_i &= -\frac{\partial H}{\partial \mathbf{q}_i} \end{aligned} \tag{1.2}$$

and $H[\{\mathbf{w}_i\}]$ is the N-particle Hamiltonian.

We go from the N-particle DF to the 1-particle DF by making the assumption that there are no correlations between particles.

$$f^{(N)}[\{\mathbf{w}_i\}] = \prod_i f(\mathbf{w}_i) \tag{1.3}$$

This is commonly interpreted to mean that f is a probabilistic distribution function from which particles are sampled: thus, the PDF.

By virtue of it being an incompressible fluid, the time evolution of the DF is better modelled mathematically: the rate of change of the DF with respect to time t, $\frac{\partial f}{\partial t}$ must negate the divergence of the DF in phase space, $\frac{\partial}{\partial \mathbf{w}} \cdot (f\dot{\mathbf{w}})$ so as to ensure the conservation of probability density in phase-space. This conservation law is a continuity equation,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\partial}{\partial \mathbf{w}} \cdot (f\dot{\mathbf{w}}) = 0 \tag{1.4}$$

that holds in general, however if $\mathbf{w}$ describes a pair of canonical coordinates, then this continuity equation is equal to a statement of Liouville's theorem, and is named the Collisionless Boltzmann Equation (CBE):

$$\frac{\partial f}{\partial t} + \frac{\partial H}{\partial \mathbf{p}} \cdot \frac{\partial f}{\partial \mathbf{x}} - \frac{\partial H}{\partial \mathbf{x}} \cdot \frac{\partial f}{\partial \mathbf{p}} = \frac{\partial f}{\partial t} + [f, H] = 0 \tag{1.5}$$

where the 1-particle Hamiltonian H is $\frac{1}{2}m\mathbf{v}^2 + \Phi(\mathbf{x})[f]$ and we have used Hamilton's equations. This injection produces a theory that describes the flow of the phase-fluid f under forces described by the potential $\Phi(\mathbf{x})$. In the context of a self-gravitating system, this potential is determined by Poisson's equation,

$$\frac{\partial^2}{\partial \mathbf{x}^2} \Phi(\mathbf{x}, t) = 4\pi\mathrm{G}M \int \mathrm{d}^3\mathbf{v} \; f(\mathbf{w}, t) \tag{1.6}$$

, where M is the total mass of the system and f is normalised to 1. One physically relevant set of canonical coordinates are the action-angle coordinates, $\mathbf{w} = (\boldsymbol{\theta}, \mathbf{J})$. The angles $\boldsymbol{\theta}$ are parameters which describe the position of a star along its trajectory, a trajectory that is then characterised by the actions $\mathbf{J}$ that are the adiabatically conserved constants of motion of the system; by definition $\mathrm{d}\mathbf{J}/\mathrm{d}t = -\partial H/\partial\boldsymbol{\theta} \equiv 0$ implies $H = H(\mathbf{J})$. This result shows the way for a (qualified) statement of quasi-stationarity amongst astrophysical systems.

The long-standing belief has been that galaxies and globular clusters are long-lived structures, with evidence such as the lack of features in the Milky Way stellar age-velocity dispersion relation (van de Sande et al. 2018) or the

observed presence of regularly rotating galaxies at high redshifts $z \sim 4.5$ with ALMA (Roman-Oliveira et al. 2023). Beyond the presence of external perturbers and mergers, galaxies are often thought to only evolve secularly. This is supported by calculations which show that the crossing time, i.e. the time taken for a star to travel the length of a system, which is also the dynamical time for linear instabilities (Weinberg 1993) within a galaxy is ~ 100 Myrs (see Binney and Tremaine 2008, Section 7.1), implying that the first linear instabilities have already long concluded, and that these systems must therefore already be in strongly perturbed configurations, away from the simple discs considered by the earliest astrophysicists (e.g. **?**).

To a first approximation, a galaxy can be modelled as being in a steady-state. This is understood in the language of distribution functions via Jean's theorem, which states that a distribution function that bears no explicit time-dependence must be a function of the constants of motion of the system: the actions, $\mathbf{J}$. This result can be seen as an extension of how the Hamiltonian must be a function of $\mathbf{J}$, since $[f(\mathbf{J}), H(\mathbf{J})] = 0$ implies $\partial f / \partial t = 0$. Astrophysical systems of interest are thus usually modelled as a sum of two interacting parts; the mean field distribution $f_0 = f_0(\mathbf{J})$ which is in a steady-state, and a small fluctuating component $f_1 = f_1(\mathbf{w}, t)$.

This near-steady-state model enables the linearisation of the CBE. The Hamiltonian can also be expressed in terms of a mean field distribution $H_0 = H_0(\mathbf{J})$ and a fluctuation potential, $\Phi_1 = \Phi_1(\mathbf{x}, t)$,

$$\frac{\partial f_0}{\partial t} + \frac{\partial f_1}{\partial t} + [f_0, H_0] + [f_1, H_0] + [f_0, \Phi_1] + [f_1, \Phi_1] = 0. \tag{1.7}$$

Assuming f_1 is small allows us to exploit the separation of scales, splitting this equation into two:

$$\frac{\partial f_0}{\partial t} + [f_0, H_0] + [f_1, \Phi_1] = 0 \tag{1.8}$$

and

$$\frac{\partial f_1}{\partial t} + [f_1, H_0] + [f_0, \Phi_1] = 0. \tag{1.9}$$

The latter equation is known as the linearised CBE (LCBE), which enables the interpretation of fluctuations as a sum of individual oscillatory/growing/damping modes. The former is often identified as a quasilinear collision operator—perhaps erroneously—for reasons I will elaborate on later.

Physicists both past and present approached dynamical problems in three broad ways: they studied equilibrium systems, for which the initial and final states were equivalent, or they studied stable systems, for which the initial state would be irrelevant because time-evolution would cause convergence onto an understood final state, or they managed to map the dynamics of the system to the harmonic oscillator, discovering the normal modes of the system.

These developments are mirrored in the study of the LCBE. Setting aside that mean-field distributions are in equilibrium, it was Landau (1946) who found that treating the LCBE as an initial value problem meant that a solution could be obtained via the Laplace transform—and that the solution for an arbitrary fluctuation on a system could be expressed in terms of a transient ballistic term and a sum of Landau modes; spatial perturbations with complex time-dependences which damped/grew in the limit of $t \to \infty$. If all the Landau modes damped, then the system would be deemed as stable—if not, then not! van Kampen (1955)-Case (1959), dissatisfied with how Landau modes were not the true eigenmodes of the system since they did not have real frequencies, showed that the normal modes of the LCBE are singular and comprise a continuum of modes. The persistence of the Case-van Kampen modes shows that while Landau's calculations appears to indicate that all fluctuations damp away spatially, they must persist in velocity-space.

What remains after the linearisation process described in equation (1.8) is a secular evolution equation: the small perturbation drives an even smaller time evolution in f_0.

Gilbert (1968; 1970), Chavanis (2012), Fouvry et al. (2015) present calculations that claim to show this secular evolution equation is equivalent to a collision operator. Noting that the time-evolution of f_0 proceeds gradually be-

cause f_1 is small, they posit that f_0's time evolution maintains the invariants of the system; i.e. $f_0 = f_0(\mathbf{J}, t)$ is always true. This permits the angle-averaging of equation (1.8), removing the angle dependency. Then, they proceed to insert the microgranulation ansatz that assumes that f_1 is sourced from the Poisson noise inherent to Poisson sampling f_0 to produce the Balescu (1960)-Lenard (1960) collision operator. This collision operator has the propensity to drive systems towards the entropy-maximising isothermal DF; that is so that $f \sim \exp(-\beta H)$—this is well-understood in the context of collisional plasmas, where systems are collisionally driven towards the Maxwellian, but does this occur for collisionless astrophysical systems?

It is just not true that quasilinear collision operators reflect the dynamics of the CBE. Not only do we not understand what is lost in the angle-averaging process, but the original derivations by Balescu and Lenard, as well as similar calculations of collision operators by Lenard and Bernstein (1958) and Rostoker and Rosenbluth (1960) show that these collision operators arise from two-particle correlations: correlations in the sampling process which frustrate the random sampling assumption underlying the connection between Liouville's theorem and the CBE. For a simpler and more modern take on the topic, Hamilton (2021) describes how the Balescu-Lenard collision operator arises from considering interactions between two stars via forces which are enhanced by the polarisation cloud as each star draws nearby stars to surround themselves. These collision operators describe the dynamics of particles and not the underlying distribution function, and so are 'red herrings' in the study of the CBE.

These techniques were found in the study of the quasi-neutral plasma; a spatially homogeneous gas of electrons superposed on a background of ions mediated by the electrostatic interaction. The electrostatic force only differs from the gravitational force in one significant way: that the force between electrons is repulsive, instead of attractive as it is for stars. This manifests

itself as a minus sign:

$$
\text{Gravitational force, } F_G = -\frac{\mathrm{G}m_1 m_2}{r^2},
$$
$$
\text{Electrostatic force, } F_E = \frac{q_1 q_2}{4\pi\epsilon_0 r^2}
$$

(1.10)

where G and ϵ_0 are the gravitational constant and the susceptibility of free space, m_1, m_2, q_1, q_2 are the masses and charges of the interacting particles, and r is the distance between them.

This simple variation in the force-law has far reaching consequences. Electrostatic repulsion between electrons leads them to spread themselves out in position space, causing spatially homogeneous systems to be energetically favoured over spatially inhomogeneous systems. Self-gravitating systems on the other hand find themselves in strongly inhomogeneous configurations, such as galaxies and globular clusters. Electrons interact via what is effectively a short-ranged force: each electron repels other electrons in their neighbourhood; causing a shielding effect first captured by Debye and Hückel (1923) that suppresses the influence that an electron has at a distance. On the other hand, each star actually gathers other stars around itself, clumping to exacerbate the influence each star exerts at a distance. The analogue of Debye shielding in gravitating systems is the Jeans (1902) instability: the gravitational potential energy of the homogeneous gas of stars exceeds its kinetic energy, and causes a large-scale collapse, heating the gas until it is in equilibrium yet again (obeys the virial theorem). An even larger difference lies between the treatment of electrostatic plasmas and self-gravitating systems, which stems from the goals of the physicists who study them. The LCBE is a perturbation theory, that describes the response of a system that is infinitesimally perturbed away from a steady-state. This specialises it towards predicting growth/damping timescales via the analysis of Landau's and producing instability criteria (via his dielectric function), but renders it untrustworthy at best when it comes to understanding the dynamics of non-linearly saturated systems. If I were a plasma physicist working to achieve fusion, my goal would be to suppress the instabilities inherent to plasmas so as to ensure magnetic confinement. Galaxies and globular

clusters, however, have already experienced their first linear instabilities, and are continuing to evolve in the non-linear regime. So what have we done to capture their dynamics?

1.3 A Brief Overview of Stellar Dynamics

When it comes to studying astrophysical bodies, dynamicists tend to fixate on two geometries: stellar spheres and axisymmetric discs; globular clusters and disc galaxies (one notes that elliptical and irregular galaxies exist too, but are neglected due to the complexity of triaxial calculations). These systems are chosen because of their symmetries: The eigenfunctions of Poisson's equation are well-understood in spherical and cylindrical coordinates—and indeed, there is a dynamical correspondence between spherical systems and the infinitely thin disc in that a spherical mean-field DF can be converted into a thin disc mean-field DF via multiplication by L_z, the angular momentum about the disc axis (Hamilton et al. 2018).

The dynamical treatment of these systems began not with spheres or discs, however: Lindblad (1927) was the first to suggest that the galactic spirals observed on what were then dubbed "spiral nebulæ" could be modelled as waves originating from Maclaurin (1801) ellipsoids in unstable equilibrium. These 'half-sectorial harmonic waves' were the linearly unstable solutions of Jeans' (1919) equation—it is worth noting that Jeans' equation (that were derived by Maxwell (1867)) focuses only on spatial perturbations, marginalising the distribution of stars in velocity space and only retaining the lowest moments.

The Jeans equations are equivalent to the incompressible Euler equations—systems that evolve according to them are therefore treated as fluids. This is in contrast to systems that evolve via the CBE, which are referred to as stellar systems.

Despite the fact that Lindblad had put forth his model decades earlier, it was not until the 1960s—after the plasma physicists had furnished most of kinetic theory—that we began to believe our long-lived astrophysical systems could be studied through the lens of linear stability. Safronov (1960) and Antonov (1961) described the linear stability analysis for thin fluid discs and spherical fluid/stellar systems respectively.

Safronov found that a fluid disc rotating rigidly can become susceptible

to an instability akin to the Jeans instability—at a critical surface density, the discs would self-partition into rings. He did not find spirals because galaxies are not rigidly rotating—to more accurately model them, we have to study differentially rotating discs: discs for which the frequency of rotation $\Omega = \Omega(r)$ is a function of distance from the centre.

This is what Toomre (1964) and Lin and Shu (1964) did. Starting from the LCBE, both studied tightly-wound spirals: spirals that oscillated many times between the centre of the disc and its furthest extent, because they realised that the rapidly oscillating over and under densities would 'cancel out', mitigating long-ranged interactions between the spiral waves—this is analogous to how the electric field of an electric multipole becomes increasingly short-ranged as the number of charges increases. The tight-winding approximation allowed them to resolve the difficulty of computing the long-ranged dynamics inherent to self-gravitating systems, expressing dynamics in terms of (spatially) local parameters.

Toomre studied local instabilities on a shearing sheet utilising Landau's theory; focusing on how small, noise-based fluctuations could be gravitationally amplified into spiral waves. In doing so he derived a criterion for the local stability of a stellar disc reminiscent of Safronov's calculation,

$$Q_s = \frac{\kappa \sigma_r}{3.36 \mathrm{G} \Sigma_0}, \tag{1.11}$$

where $\kappa = \sqrt{4r\frac{\partial \Omega^2}{\partial r} + \Omega^2}$ is the epicycle frequency, σ_r is the radial velocity dispersion, G is the gravitational constant, and Σ_0 is the surface density of the thin disc. It can be intuited that Toomre's Q is a measure of how 'hot' the disc is: with a kinetic energy numerator, and a gravitational potential energy denominator. When $Q > 1$, the disc is hot and thus stable. When $Q < 1$, the disc becomes increasingly cold, feeding the same mechanism that leads to the Jeans instability.

While Toomre provided a mechanism that could give birth to a spiral, Lin & Shu focused on deriving a theory of global standing waves on a disc galaxy, so as to directly capture spirals—for this reason, the theory is known

as "Quasi-stationary density wave theory". For a wave with the potential $\Phi_1 = \Re[A \exp(\mathrm{i}(\omega t - m\phi - kr)]$, they produced the Lin-Shu (L-S) dispersion relation for a fluid disc:

$$(\omega - m\Omega)^2 = c_s^2 k^2 + \kappa^2 - 2\pi \mathrm{G}\Sigma_0 |k| \tag{1.12}$$

Where c_s is the speed of sound in the fluid, a surrogate for the radial velocity dispersion found in the stellar disc. The condition for instability (complex-valued ω) could be obtained by setting the RHS of this equation to zero: doing so produces the fluid disc Toomre Q;

$$Q_f = \frac{\kappa c_s}{\pi \mathrm{G}\Sigma_0} \tag{1.13}$$

that plays the same role as a stability criterion. This theory describing a fluid disc was extended by Lau and Bertin (1978) to accomodate for how the fluid would move within the potential of a (stellar) spiral arm, resulting in global mode theory.

To describe spirals, however, it is not enough that we explain how some arbitrary waves arise in a disc: we must describe how they have survived until today.

Toomre (1969) showed that L-S waves were not stationary, but rather carried energy with them with the group velocity $v_g = \frac{\partial \omega}{\partial k}$ radially. He found that in our solar neighbourhood, $v_g \approx 12.5 \mathrm{kpc/Gyr}$—a spiral arm like the Perseus arm would have 'left' the Milky Way in that time. This meant that these spiral waves had to be regenerated in some way.

There are two theories for how this could have happened. The first is swing amplification (Goldreich and Lynden-Bell 1965, Julian and Toomre 1966): Stars moving in orbits slightly perturbed from circular orbits can be approximated by stars moving in elliptical orbits. Noise within the disc seeds equal amounts of leading and trailing spirals in the noise, but as the galaxy shears, the small leading spirals begin shearing into trailing spirals. In this process, they become radial over-densities for a moment, co-rotating with the

elliptical orbits of neighbouring stars and thus drawing them in to bolster itself. Swing amplification is thought to seed chaotic, flocculent spirals since it is sourced from noise/Giant Molecular Cloud perturbers (Sellwood 2011). However, numerical studies (Grand et al. 2013, D'Onghia 2015) provide rationale for how one can extend swing amplification theory that focuses on local instabilities to obtaining the most strongly amplified global m-modes.

The second is the 'groove instability'. Sellwood and Kahn (1991), Sellwood and Carlberg (2019), Sellwood and Carlberg (2022) propose that narrow but steep deficiencies in angular momentum (i.e. grooves cut into the distribution function at a fixed angular momentum) amplify initially small, infinitesimal perturbations until non-linear saturation is achieved in the form of a spiral. The spiral then serves to transport angular momentum (e.g. what is known as radial migration) within the system, filling in the old groove and seeding new grooves on the outer Lindblad resonance of the spiral, to restart the cycle. The groove instability is a far weaker instability than swing amplification is; as a result the former 'piggy-backs' on the latter.

Both these methods, however, fail to replicate the grand design $m = 2, 4$ spirals that we observe in reality. Swing amplification is sourced from random noise and thus does not produce symmetric spirals.

The groove instability seeding a 'grand design spiral' as Sellwood and Carlberg (2022) demonstrate requires either allowing only the $m = 2$ component of their gravitational potential run in their restricted simulation (effectively ensuring that discs with two spiral arms are born) or requiring that the $m = 2$ seed amplitude is boosted (effectively ensuring that the features with two spiral arms become dominant first before other m-numbers set in). They also show that the newly reseeded groove instabilities interfere with the declining nonlinearly saturated spiral that already exists in a complex way.

I have just loosely described mechanisms how initially axisymmetric discs may 'spontaneously' produce spirals. These are not the only mechanisms that have been proposed to form spirals, however. Barred galaxies have been pro-

posed to 'leak' stars at their tips (unstable Lagrange points) that form spirals that co-move with the bar's resonance manifolds: (Athanassoula et al. 2009a;b; 2010). Tagger et al. (1987) suggested that bars may nonlinearly couple with spirals (see Masset and Tagger 1997, for simulations), allowing the former to transfer energy to the latter. Finally, tidal interactions between galaxies have been attributed with producing grand design galaxies, an idea owed to Holmberg (1941), who made physical simulations of galaxies with lightbulbs and photodetectors, that is mainly supported by simulations (early on, see Toomre and Toomre (1972) and later, see Dobbs et al. (2010)).

While the dynamical approach has been successful in describing how spirals arise, it is also limited because these theories utilise tools that are adapted to use in the context of plasmas: which is best poised to explore linear instabilities, not ones that develop non-linearly. In general, the mechanisms that begin with an axisymmetric disc are able to carefully describe how a nascent spiral is born, but then falter when it comes to describing the dynamics of the spiral as it persists. Similarly, calculations that support bar-driven spirals and tidal interactions do not explain why galaxies are susceptible to spiral perturbations—they only present that they are forced by these external perturbers.

Globular clusters are in a collisional regime, and so Chandrasekhar (1949) scattering is believed (e.g. see Hamilton et al. 2018, Lau and Binney 2019, Heggie et al. 2020, Lau and Binney 2021c, for some discussion on whether or not this is the case) to describe their secular evolution. A linear instability of globular clusters known as the radial orbit instability was first uncovered by Antonov (1973) and then corroborated by Henon (1973) and later Polyachenko and Shukhman (1981) in simulations. This instability occurs when a spherical cluster is sufficiently anisotropic in velocity space such that the orbits comprising the cluster are sufficiently eccentric. Then these orbits may undergo spontaneous alignment, producing an (as it was in the theory of spiral galaxies) poorly understood triaxial equilibrium.

What we need is a theory of correlations: a theory that describes the natural correlations between two different points in phase-space of the CBE. This theory would not tell which of the 'modes' of a system would grow to dominate the system most quickly, instead describing how each star influences the likely positions of other nearby stars—this is a statistical mechanical theory.

Traditional treatments of stellar systems (i.e. the usage of Jeans' equation) were expedient in the past because they described the dynamics of the quantities that were then observed: the velocity dispersion of stars, their streaming velocity, and their density. As simulations and observations became more and more nuanced, treatments based on the CBE became more and more ubiquitous, reflecting greater confidence in our knowledge of velocity-space observations.

We have now reached the point where our observational capabilities have far superseded those of our plasma physicist counterparts in terms of our ability to resolve individual particles. It is simply impossible to track the coordinates of a single electron or ion in a plasma, but we can observe parallaxes and proper motions for stars to a high precision, and this has only become more and more true in the Gaia era (see Gaia Collaboration et al. 2022).

On the other hand, there is much we do not understand about the bulk dynamics of galaxies. The presence of dark matter not only means that we do not know what the initial condition should be since we only see the baryonic matter, but also means that we do not fully understand the dynamical friction between baryonic and dark matter. If we recall, Maxwell's Maxwellian was first derived solely on the grounds of energy being a collisional invariant of gases—so we cannot make arguments on that ground either, because we cannot see what happens to approximately half of our 'gas'! There is no better proof of this than the ever-growing list of mean field distribution functions f_0 that we astrophysicists hand-pick to model simulations and observations. Statistical mechanics gives us the ability to compute f_0 given an understanding of a system's underlying dynamics, and vice versa, and is thus well suited to

cutting through the practice of choosing distribution functions so as to fit an observation.

Connecting f_0 to the fluctuations we expect to see on a system is solely the purview of statistical mechanics.

1.4   Dynamics & Statistical Mechanics

The advantages of treating systems as fluids instead of massive collections of particles have made it common practice to model our microstates via the phase-space density f.

At first glance, f seems to be a suitable candidate for a macroscopic quantity: obtaining the DF allows us to marginalise the microstates, and many different microstates can be Poisson sampled from the same distribution function. We cannot assign f to this role, however, since many DFs can be Poisson sampled to produce the same microstate too.

This many-to-many relationship between the DF and the microstate is not unexpected. The DF is just a model for the dynamics of a microstate, and there are many models that can describe the same dynamics. We defined macroscopic features as the features that are common between all representative models of a system. We also know that the macrostate can be represented by the distribution of representative models of a system—thus it is clear that the macroscopic features of a system can be described by the moments of the distribution of representative models.

This generalisation of the definition of a macroscopic feature allows us to unify dynamics and statistical mechanics—whereas prior to this, statistical mechanics—even non-equilibrium statistical mechanics—and dynamics have been treated as two separate fields.

To understand why this is so, we must look into modern statistical mechanics, which can be roughly segmented into two pursuits: that of non-extensive statistical mechanics and non-equilibrium statistical mechanics.

Extensivity as a concept in physics begins and ends with Boltzmann-Gibbs statistical mechanics (BGSM). If a macroscopic quantity is extensive, then it scales linearly with system-size. 'System-size' can be loosely thought of as the number of particles within the system: it describes the size of the configuration-space; the total number of microstates. For the BGSM to hold, both the entropy and the energy has to be extensive: well the former might

be, but the latter simply is not for astrophysical systems!

Non-extensive statistical mechanics, reviewed by Tsallis (2011) but also see his original paper Tsallis (1988), is based on a simple premise: for BGSM to hold, a system's N constituent particles have to be uncorrelated from each other.

This implies the multiplicative nature of probability; the probability of particles being observed in a configuration defined by the set $\{\mathbf{w}_i\}$ is thus $\prod_i p(\mathbf{w}_i)$, the product of individual probabilities evaluated at the positions of particles p_i. The expectation of this probability of observation is $\prod_i p(\mathbf{w}_i) = \exp(\sum_i \ln p(\mathbf{w}_i)) \rightarrow \exp(\sum_\Delta p_\Delta \ln p_\Delta) = \exp(-S)$ which defines the Shannon entropy, S as a sum (or integral) of $-p \ln p$ over all phase-space elements Δ.

If the system is interacting such that the particles are no longer uncorrelated, then perhaps the Shannon entropy is no longer suitable: a non-additive entropy can be chosen to ensure the extensivity of the system, and an extension of BGSM is made by choosing a 'generalised entropy'.

Generalised entropies are concave, continuous functions (Khinchin 1957) of p_i, the occupancy number of a state indexed by i and have been referred to within astrophysics as H-functions (see Tremaine et al. 1986). They are exemplified by the Tsallis (T_q) and Renyi (R_q) entropies but of course include the entropy of Shannon's (S).

$$T_q = \frac{1 - \sum_i p_i^q}{1 - q}$$
$$R_q = \frac{1}{1 - q} \ln \left(\sum_i p_i^q \right) \tag{1.14}$$
$$S = -\sum_i p_i \ln p_i$$

The usage of H-functions in the astrophysical context are a direct consequence of applying the principle of maximum entropy (a cornerstone of BGSM) to derive the mean field distribution functions found in astrophysical contexts. The Shannon entropy can only produce isothermal distributions, after all, while the Tsallis distributions produce distributions with tails that are power-law in the energy.

This theory is not suitable for astrophysical macroscopic observables (the reliance on the connection between time averages and ensemble averages ensures this!). However this is an issue with the Boltzmann-Gibbs formalism, and not the generalised entropies: recall that the entropy for a non-interacting boson/fermion gas is

$$S_{b/f} = \sum_i p_i \ln p_i + \begin{cases} \text{if boson; } \sum_i (1 + p_i) \ln(1 + p_i) \\ \text{if fermion; } - \sum_i (1 - p_i) \ln(1 - p_i) \end{cases}, \qquad (1.15)$$

where the sums run over all distinguishable states indexed by i, and p_i is the occupancy number, the average number of particles occupying the i-th state. Even though these particles are non-interacting, their entropies describe the qualities of these particles. In the case of a fermion, it is clear that the additional term describes Pauli's exclusion principle: each state can only be occupied by one fermion, thus the additional $-(1 - p_i) \ln(1 - p_i)$ contribution. In the case of the boson, we note that p_i is the average occupancy, not the probability of a single sampling—for the sake of calculation, let us denote the probability of a single sample as x. Then a boson gas is populated by allowing more than one particle to be populated per sample, causing solving for p_i in terms of x to reduce to solving a geometric progression,

$$p_i(x) = \frac{\sum_{n=0}^{\infty} n x^n}{\sum_{n=0}^{\infty} x^n} = \frac{1}{1/x - 1}$$

$$S_b(x) = -\frac{\sum_{n=0}^{\infty} x^n \ln(x^n)}{\sum_{n=0}^{\infty} x^n} = -\frac{x}{1 - x} \ln x - \ln(1 - x)$$

$$(1.16)$$

Substituting x for p_i produces the entropy for a non-interacting boson gas in terms of its average occupancy in agreement with equation (1.15). Since p_i is a monotonically increasing function for $0 \leq x < 1$, it may not seem like it matters whether we express the entropy with respect to p_i or x. Notice that p_i is what we use to take moments of the energy, since it is a descriptor for the 'average microstate' and thus plays the role of the model that f does for stars.

The entropy incorporates dynamical information relating the density of particles with the probability that they are realised, while the energy mediates

interactions between densities of particles in a manner that is insensitive to the distribution of particles underlying the densities. Generalised entropies become relevant when the density of particles (once normalised) no longer coincides with the probability distribution of particles.

Non-equilibrium statistical mechanics, on the other hand, focuses on describing systems that have been driven out of equilibrium—and are usually 'near-equilibrium'. It begins with the combination of phenomenological laws describing resistances (Newton's law of viscosity or Ohm's law), their combination with a fluid conservation equation for mass/charge (the Euler equations or Kirchoff's law) to produce a theory like the Navier (1821)-Stokes (1845)/Burgers (1948) equation or the Chapman-Enskog (Chapman and Cowling 1990) equations.

Then tools that are meant to describe how these resistances function are introduced via the language of stochastic calculus. This is done by characterising the surroundings of the particle undergoing this resistance as a heat-bath with 'baked-in' fluctuations, as when Einstein (1905) and von Smoluchowski (1906) characterised the diffusion of particles undergoing Brownian motion.

However, the connection between the heat bath and the particles was only fully formalised by the Langevin et al. (1908) equation,

$$m\ddot{q} = -V'(q) + \int_0^t \mathrm{d}s \; D(t-s)\dot{q}(s) + \eta(t) \tag{1.17}$$

describes the acceleration of a particle with position q undergoing Brownian motion within a potential $V(q)$: it is decelerated by the second term on the RHS, that handles the dynamical friction on the particle by modelling it via a deterministic autocorrelation $D(t)$, and is kicked about by the stochastic force $\eta(t)$ that represents the noise inherent to the system.

This creates a genuinely dissipative process: the relationship between $D(t)$ and η is known as the fluctuation-dissipation theorem (Green 1954, Kubo 1957).

$$\langle \eta(t)\eta(t') \rangle = \frac{1}{\beta}\dot{D}(t-t'), \tag{1.18}$$

where β is the temperature of the system and assuming that η is governed by a stationary Gaussian process: the ensemble average brackets $\langle\rangle$ here denote an averaging over that random process.

Non-equilibrium statistical mechanics at this degree of approximation describes how a single particle jitters about in a heat bath that is far larger than itself, so that the heat bath is not perturbed by its interactions with the particle, while the particle's autocorrelation in time is reduced dissipatively. This does not align with our study of (nearly) collisionless astrophysical systems, for which true dissipation is not thought to exist (see Section 1.2 of Binney and Tremaine 2008).

Non-equilibrium statistical mechanics also falls short of unifying dynamics and statistical mechanics because it fixes the heat bath: in the astrophysical context, the approximations involved are similar to integrating each star as a massless tracer within the potential of a field of stars: while this gives us insight into the structure of the 'heat bath' by telling us how quickly the massless tracer loses the information it initially had, we exclude the response of the heat-bath—and that is essentially what happens with the CBE, where f flows under its own Hamiltonian.

In this thesis I go from studying conserved quantities of the linearised CBE to formulating a generalisation of Boltzmann-Gibbs statistical mechanics that utilises Shannon's definition of the entropy.

Chapter 2 (Lau and Binney 2021a) elucidates the conserved quantities of fluctuations on isotropic spherical systems. We find that by dividing the perturbation into even and odd parts, we can convert the first-order linearised CBE into a pair of second-order wave equations. Their solutions, the 'quadratic van Kampen modes' reveal the structure of the solutions to the LCBE: not only is the energy of the perturbation (Binney and Tremaine 2008, equation 5.130) positive-semidefinite, but also linearly unstable modes are purely growing/damping and do not contribute to the energy while the stable modes are oscillatory and do.

Chapter 3 (Lau and Binney 2021b) generalises this expression for the energy of a linear fluctuation to non-ergodic systems, noting that the energy could be either positive or negative without issue. We show that the orthodox linear van Kampen modes conserve this energy too. It was at this point when I began to realise that while the van Kampen modes could in principle describe any linear fluctuation, they could not tell us which fluctuation are more likely than others.

Chapter 4 (Lau and Binney 2021d) describes how we can convert the Collisionless Boltzmann Equation from an algebraic ODE to a functional ODE. While it is mathematically simple to do—replace position and velocity derivatives with position and velocity derivatives of functional derivatives, it represents a change in the way we understood macroscopic quantities—from phase-space averages, time-averages, angle-averages or fast-action averages, to functional averages with a measure over the space of all possible f.

Chapter 5 (Lau 2023) then describes how we can compute macroscopic quantities under this new generalised theory—how we might define the distribution of representative models of a system, and what goes into a model that represents the dynamics of a system. This allows us to define relationships between macroscopic quantities much like traditional thermodynamics does.

Chapter 2

Modes of a Stellar System I: ergodic systems[1]

2.1   Abstract

The excursions of star clusters and galaxies around statistical equilibria are studied. For a stable ergodic model Antonov's Hermitian operator on six-dimensional phase space has the normal modes as its eigenfunctions. The excitation energy of the system is just the sum of the (positive) energies associated with each normal mode. Formulae are given for the DFs of modes, which are of the type first described by van Kampen rather than Landau, and Landau 'modes' can be expressed as sums of van Kampen modes. Each van Kampen mode comprises the response of non-resonant stars to driving by the gravitational field of stars on a group of resonant tori, so its structure is sensitive to the degree of self gravity. The emergence of global distortions in N-body models when particles are started from an analytical equilibrium is explained in terms of the interplay of normal modes. The positivity of modal energies opens the way to modelling the thermal properties of clusters in close analogy with those of crystals.

---

[1]This chapter is taken from Lau and Binney (2021a).

2.2 Introduction

Galaxies and star clusters are in approximate states of equilibrium and have for decades been fitted to models in which the distribution function $f(\mathbf{x}, \mathbf{v})$ of their constituent particles (stars, dark-matter particles) are steady-state solutions of the collisionless Boltzmann equation (CBE). The advent of massive simulations of galaxy formation (Laporte et al. 2019) and detailed data from the Gaia mission (Gaia Collaboration and Brown 2018) and large integral field units such as MUSE (e.g. Vitral and Mamon 2021) have stimulated interest in non-equilibrium features of galaxies, especially the Milky Way (Antoja et al. 2018).

For almost a century observations of galaxies and star clusters have been interpreted in terms of 'mean-field' models, that is to say models in which fluctuations have been averaged away. In the case of globular clusters the community has been aware since at least the pioneering work of Hénon (1961) that fluctuations drive secular evolution of the system towards higher central concentration and lower mass (core collapse and evaporation) but observations have nonetheless been fitted to mean-field models on the grounds that clusters evolve through a series of mean-field models.

Fluctuations in the surface brightnesses of early-type galaxies form the basis for a standard technique for estimating their distances (Tonry and Schneider 1988), but the fluctuations are computed by imposing shot noise on equilibrium models rather than using a dynamical theory of fluctuations.

Perhaps the most exciting single discovery made in the Gaia DR2 data is the phase spiral that Antoja et al. (2018) uncovered in the distribution of stars in the (z, v_z) plane. The spiral is surely a symptom of a macroscopic oscillation of the disc that has a significant component in the z direction. If we had a credible dynamical model of this oscillation, we would be able to extract from the Gaia data information about the structure of the disc and the agent [likely the Sagittarius dwarf galaxy (Binney and Schönrich 2018, Laporte et al. 2019, Bland-Hawthorn and Tepper-Garcia 2020)] that excited it.

An impediment to this program is how the the disc's self-gravity certainly plays an important role in the oscillation, and there is little prospect of adequately modelling the disc's oscillation until we have a better understanding of the global oscillations of stellar systems. This is the first in a series of thesis chapters that lay the foundations for such understanding by setting up an adequate theory of the normal modes of stellar systems.

Normal modes (in quantum mechanics 'stationary states') owe their usefulness to three key properties: (i) they are complete in the sense that any initial condition can be expressed as a linear combination of normal modes; (ii) they have the trivial time dependence $e^{-i\omega t}$; (iii) they are mutually orthogonal, with the consequence that the energy of the whole system is simply the sum of the energies invested in each normal mode.

Modes of stellar systems have received significant attention since the work of Toomre (1964), Lin and Shu (1964), and Kalnajs (1965). That work was motivated by the desire to understand spiral structure so focused on 'razor-thin' rotating stellar discs which were confined to evolving in a plane. Two decades later the focus switched to hot, spherical systems from a desire to understand how and when radial bias in the velocity dispersion caused systems to lose spherical symmetry (Palmer and Papaloizou 1987, Saha 1991, Weinberg 1991). The standard reference for this work is the two volumes of Fridman and Polyachenko (1984), and a glance at the contents pages make clear that interest focused exclusively on the search for unstable normal modes. We show below that these modes are qualitatively different from the modes required to investigate, as we do, the excursions that stable systems make around equilibrium.

Fluctuations may be externally or internally driven. The Antoja spiral in our Galaxy and shells around early-type systems (Malin and Carter 1980) are surely externally driven. The secular evolution of globular clusters is largely driven by fluctuations that are internally driven by Poisson noise [although fluctuations driven externally by tidal fields are also significant (Lee and Os-

triker 1987)]. Even after more than a half century of work, there is no consensus as to whether observed spiral structure is sometimes internally driven (Sellwood & Masters 2022), although some 'grand-design' spiral structure (e.g., that of M51) is certainly externally driven. Whatever the driving mechanism, we choose to model fluctuations as solutions to the linearised Boltzmann equation (CBE) coupled to the already linear Poisson equation.

This time-translation invariant pair of linear equations may be expected to have a complete set of solutions with time dependence $\mathrm{e}^{-\mathrm{i}\omega t}$ (with potentially complex ω). In this chapter we derive these solutions for the important special case that the unperturbed system is ergodic – that is has a distribution function (DF) of the form $f_0(H)$, where

$$H(\mathbf{x}, \mathbf{v}) = \tfrac{1}{2}v^2 + \Phi(\mathbf{x}) \tag{2.1}$$

is the Hamiltonian of a single particle moving in the gravitational potential Φ. The second contribution (Chapter 3) in the series generalises many results to the case of a DF of the form $f_0(\mathbf{J})$, where $\mathbf{J}$ is the vector of the action integrals of stars moving in the unperturbed potential. In the presently unpublished third contribution in the series we develop an apparatus for decomposing an arbitrary initial condition of a system with $f_0(\mathbf{J})$ into its constituent normal modes. The present paper relies heavily on an Hermitian operator that Antonov (1961) introduced. This operator does not generalise straightforwardly from ergodic systems to more general ones, so Chapter 3 obtains a restricted range of results with a simpler but less powerful technique. The fourth unfurnished contribution in the series generalises Antonov's operator to DFs of the form $f(\mathbf{J})$.

The plan of this chapter is as follows. Section 2.3 introduces basic concepts and establishes notation. Section 2.4 introduces the Hermitian operator K on phase space whose eigenfunctions are the required normal modes of the cluster. If the system is stable, all its modes are van Kampen modes; they have real frequencies drawn form a continuous spectrum. If the system is unstable the spectrum contains isolated pure imaginary frequencies. We show

that the energies of modes are additive, and give a very simple expression for a mode's energy in terms of its DF. This expression implies that the energy of van Kampen modes is positive and that of modes with imaginary frequencies vanishes. We show also that K gives rise to a slightly different conserved quantity that provides a means to establish stability. In Section 2.5.1 we show that K commutes with the angular-momentum operator L_z before in Section 2.5.2 obtaining an expression for the DF of a van Kampen mode. This contains a free function and parameters that can be computed from the free function by matrix algebra. In Section 2.5.4 we investigate the way in which the structures of a van Kampen mode depends on the extent to which a system is self-gravitating. In Section 2.5.5 we re-express a van Kampen mode's energy in terms of the free function and the potential that the mode generates, and in Section 2.5.6 we discuss the emergence of system-scale fluctuations in N-body simulations. In Section 2.6 we argue that the van Kampen modes are complete and do not result in singular distribution functions. In Section 2.7 we stress the importance of the concept of particle dressing in stellar dynamics as in other branches of physics, and discuss the role that van Kampen modes play in dressing. Section 2.7.2 discusses the relationship between van Kampen and Landau modes, while Section 2.7.3 considers the prospect for using van Kampen modes to extend conventional statistical mechanics to stellar systems, and for understanding the role of thermal fluctuations within them. Section 2.8 concludes.

2.3   Mathematical background

Here we introduce essential mathematical tools and establish our notation. We focus on stable ergodic clusters, that is systems with unperturbed DFs $f_0(H)$ where

$$H(\mathbf{x}, \mathbf{v}) \equiv \tfrac{1}{2}v^2 + \Phi(\mathbf{x}) \tag{2.2}$$

is the Hamiltonian of a single particle in the gravitational potential $\Phi(\mathbf{x})$. A necessary and sufficient condition for such a system to be stable is that the

derivative $f_0' < 0$ at all energies (Antonov 1961).

2.3.1 Variable degree of self gravity

In the following it proves helpful to be able to consider self-gravity to be a variable ξ that runs from zero (stars move in the fixed potential of a specified density distribution) to unity (stars experience only their gravitational attraction to the other stars). It is straightforward to set up a simulation for any given value of ξ by sampling an analytic density distribution in the usual way and taking the force on each star to be ξ times the force returned by an N-body solver plus $(1 - \xi)$ times the force provided by the analytic density.

2.3.2 Angle-action variables

The role that Cartesian variables play for homogeneous systems is played for spheroidal systems by angle-action variables $(\boldsymbol{\theta}, \mathbf{J})$. The actions J_i are constants of motion while their conjugate variables, the angles θ_i, increase linearly in time, so $\boldsymbol{\theta}(t) = \boldsymbol{\theta}(0) + \boldsymbol{\Omega}t$. The particles' Hamiltonian $H(\mathbf{x}, \mathbf{v})$ is a function $H(\mathbf{J})$ of the actions only and the frequencies Ω_i that control the rates of increase of the angles are given by $\boldsymbol{\Omega} = \partial H / \partial \mathbf{J}$. Angle-action variables are canonical, so the volume element of phase space $\mathrm{d}^6\mathbf{w} = \mathrm{d}^3\mathbf{x}\mathrm{d}^3\mathbf{v} = \mathrm{d}^3\boldsymbol{\theta}\mathrm{d}^3\mathbf{J}$ and Poisson brackets can be computed as

$$[f, g] = \sum_i \left(\frac{\partial f}{\partial \theta_i} \frac{\partial g}{\partial J_i} - \frac{\partial f}{\partial J_i} \frac{\partial g}{\partial \theta_i} \right). \tag{2.3}$$

Functions on phase space can be expressed as Fourier series:

$$h(\mathbf{w}) = \sum_{\mathbf{n}} h_{\mathbf{n}}(\mathbf{J}) \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}} \ ; \ h_{\mathbf{n}}(\mathbf{J}) = \int \frac{\mathrm{d}^3\boldsymbol{\theta}}{(2\pi)^3} \mathrm{e}^{-\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}} h(\mathbf{w}). \tag{2.4}$$

Note that for real h, $h_{-\mathbf{n}} = h_{\mathbf{n}}^*$.

2.3.3 Potential-density pairs

While the potential $\Phi(\mathbf{x})$ is a function of only $\mathbf{x}$, it becomes a function of both $\boldsymbol{\theta}$ and $\mathbf{J}$. So while angle-action variables simplifies dynamics into a matter of evolving $\boldsymbol{\theta}$ linearly in time, they seriously complicate the solution of Poisson's equation. Following Kalnajs (1976) this difficulty is finessed by introducing a

basis of biorthogonal potential-density pairs. That is, a set of pairs $(\rho^{(\alpha)}, \Phi^{(\alpha)})$ such that

$$4\pi G \rho^{(\alpha)} = \nabla^2 \Phi^{(\alpha)} \quad \text{and} \quad \int \mathrm{d}^3\mathbf{x}\, \Phi^{(\alpha)*} \rho^{(\alpha')} = -\mathcal{E} \delta_{\alpha\alpha'}, \tag{2.5}$$

where $\mathcal{E}$ is an arbitrary constant with the dimensions of energy. Given a density distribution $\rho(\mathbf{x})$, we expand it in the basis

$$\rho(\mathbf{x}) = \sum_\alpha A_\alpha \rho^{(\alpha)}(\mathbf{x}) \quad \Leftrightarrow \quad \Phi(\mathbf{x}) = \xi \sum_\alpha A_\alpha \Phi^{(\alpha)}(\mathbf{x}), \tag{2.6}$$

where

$$A_\alpha = -\frac{1}{\mathcal{E}} \int \mathrm{d}^6\mathbf{w}\, \Phi^{(\alpha)*}(\mathbf{x}) f(\mathbf{w}). \tag{2.7}$$

If ρ and Φ are time-dependent, the A_α become time-dependent. From equations (2.5) and (2.6) one can obtain an expression for Φ in terms of ρ. Comparison of this relation with Poisson's integral, yields

$$\frac{G}{|\mathbf{x}' - \mathbf{x}|} = \frac{1}{\mathcal{E}} \sum_\alpha \Phi^{(\alpha)}(\mathbf{x}) \Phi^{(\alpha)*}(\mathbf{x}'). \tag{2.8}$$

2.4   Antonov's operator K

We now derive for stable ergodic clusters the Hermitian operator K introduced by Antonov (1961). The true normal modes of the system are eigenfunctions of K with non-negative eigenvalues that prove to be the squares of the modes' (real) frequencies. Following Antonov (1961) we split the perturbed DF f_1 into parts that are even and odd in $\mathbf{v}$:

$$f_1(\mathbf{x}, \mathbf{v}) = f_+(\mathbf{x}, \mathbf{v}) + f_-(\mathbf{x}, \mathbf{v}) \tag{2.9}$$

where

$$f_\pm(\mathbf{x}, \mathbf{v}) \equiv \tfrac{1}{2}[f_1(\mathbf{x}, \mathbf{v}) \pm f_1(\mathbf{x}, -\mathbf{v})]. \tag{2.10}$$

In the absence of a perturbation, f_- vanishes, so this part of the DF isolates the effect of the perturbation. On the other hand the perturbation changes the potential only through f_+.[2]

---

[2]In Dirac's seminal textbook, he argues that the second-order Klein-Gordon equation cannot stand in for the Schrödinger equation, which is first-order in time. So he factorises the Klein-Gordon operator into two first-order operators, by splitting the wavefunction into four parts. Antonov proceeded in the opposite direction: by splitting the DF he derived two first-order operators and then combined them to obtain a second-order equation.

We define an inner product on the space of DFs by

$$\langle g|f\rangle \equiv -\int d^6\mathbf{w}\, \frac{g^*f}{f_0'(H)}, \tag{2.11}$$

where $H(\mathbf{J})$ is the unperturbed Hamiltonian and the leading minus reflects the fact that $f_0' < 0$. Since $d^6\mathbf{w}\, f$ has dimensions of mass and f/f_0' has dimensions of v^2, $\langle g|f\rangle$ has the dimensions of Mv^2, i.e., energy. When we Fourier expand the DFs we find

$$\langle g|f\rangle = -(2\pi)^3 \int \frac{d^3\mathbf{J}}{f_0'} \sum_{\mathbf{k}} g_{\mathbf{k}}^* f_{\mathbf{k}}. \tag{2.12}$$

Notice that

$$\langle f_1|f_1\rangle = \langle f_-|f_-\rangle + \langle f_+|f_+\rangle. \tag{2.13}$$

$H(\mathbf{J})$ is even in $\mathbf{v}$ and the Poisson bracket operator is odd in $\mathbf{v}$, so our division (2.9) of f splits the linearised CBE, $\partial_t f_1 + [f_1, H] + [f_0, \Phi_1] = 0$, into two equations

$$\frac{\partial f_+}{\partial t} = -[f_-, H] \quad ; \quad \frac{\partial f_-}{\partial t} = -[f_+, H] + [\Phi_1, f_0]. \tag{2.14}$$

Now $[\Phi_1, f_0] = f_0'(H)[\Phi_1, H]$, so

$$\begin{aligned}
\frac{\partial}{\partial t}[\Phi_1, f_0](\mathbf{w}) &= f_0'(H)\left[\frac{\partial \Phi_1}{\partial t}, H\right] \\
&= -f_0'(H)\left[\int d^6\mathbf{w}'\, \frac{\xi G}{|\mathbf{x}' - \mathbf{x}|}\frac{\partial f_+(\mathbf{w}')}{\partial t}, H(\mathbf{w})\right].
\end{aligned} \tag{2.15}$$

We differentiate the second of equations (2.14) wrt t and use the first equation to eliminate $\partial_t f_+$ from the rhs to obtain

$$\begin{aligned}
\frac{\partial^2 f_-}{\partial t^2} &= [[f_-, H], H] \\
&+ f_0'(H)\left[\int d^6\mathbf{w}'\, \frac{\xi G}{|\mathbf{x}' - \mathbf{x}|}[f_-(\mathbf{w}'), H(\mathbf{w}')], H(\mathbf{w})\right].
\end{aligned} \tag{2.16}$$

This equation is of the form

$$\frac{\partial^2 f_-}{\partial t^2} = -Kf_-, \tag{2.17}$$

where the operator

$$
\begin{aligned}
Kf_- &\equiv -[[f_-, H], H] \\
&- f_0'(H)\left[\int \mathrm{d}^6\mathbf{w}' \frac{\xi G}{|\mathbf{x}' - \mathbf{x}|}[f_-(\mathbf{w}'), H(\mathbf{w}')], H(\mathbf{w})\right].
\end{aligned}
\tag{2.18}
$$

In terms of angle-action coordinates $(\boldsymbol{\theta}, \mathbf{J})$,

$$
[f(\boldsymbol{\theta}, \mathbf{J}), H(\mathbf{J})] = \boldsymbol{\Omega} \cdot \frac{\partial f}{\partial \boldsymbol{\theta}},
\tag{2.19}
$$

so K can be written

$$
\begin{aligned}
Kf_- &= -\left(\boldsymbol{\Omega} \cdot \frac{\partial}{\partial \boldsymbol{\theta}}\right)^2 f_- \\
&- f_0'(H)\,\boldsymbol{\Omega} \cdot \frac{\partial}{\partial \boldsymbol{\theta}} \int \mathrm{d}^6\mathbf{w}' \frac{\xi G}{|\mathbf{x}' - \mathbf{x}|}\boldsymbol{\Omega}' \cdot \frac{\partial f_-(\mathbf{w}')}{\partial \boldsymbol{\theta}'}.
\end{aligned}
\tag{2.20}
$$

Inserting equation (2.8) into (2.20) yields

$$
\begin{aligned}
Kf_- &= -\left(\boldsymbol{\Omega} \cdot \frac{\partial}{\partial \boldsymbol{\theta}}\right)^2 f_- - f_0'(H)\frac{\xi}{\mathcal{E}}\boldsymbol{\Omega} \cdot \frac{\partial}{\partial \boldsymbol{\theta}} \\
&\times \int \mathrm{d}^6\mathbf{w}' \sum_\alpha \Phi^{(\alpha)}(\mathbf{x})\Phi^{(\alpha)*}(\mathbf{x}')\boldsymbol{\Omega}' \cdot \frac{\partial f_-(\mathbf{w}')}{\partial \boldsymbol{\theta}'}.
\end{aligned}
\tag{2.21}
$$

At this point it's convenient to define

$$
j_\alpha[f_1](t) \equiv -\frac{\mathrm{i}}{\mathcal{E}} \int \mathrm{d}^6\mathbf{w}\, \Phi^{(\alpha)*}(\mathbf{x})\boldsymbol{\Omega} \cdot \frac{\partial f_-(\mathbf{w})}{\partial \boldsymbol{\theta}},
\tag{2.22}
$$

because it allows us to write equation (2.21) in the form

$$
K = -\left(\boldsymbol{\Omega} \cdot \frac{\partial}{\partial \boldsymbol{\theta}}\right)^2 f_- - \mathrm{i}\xi f_0'(H)\,\boldsymbol{\Omega} \cdot \frac{\partial}{\partial \boldsymbol{\theta}} \sum_\alpha \Phi^{(\alpha)}(\mathbf{x})j_\alpha[f_-].
\tag{2.23}
$$

j_α is a functional of f_1 rather than just f_- because it can also be computed from f_+: eliminating f_- using the first of equations (2.14) and equation (2.19) we obtain

$$
j_\alpha[f_1](t) = \mathrm{i}\frac{1}{\mathcal{E}} \int \mathrm{d}^6\mathbf{w}\, \Phi^{(\alpha)*}(\mathbf{x})\frac{\partial f_+}{\partial t}.
\tag{2.24}
$$

Since by equation (2.7) the coefficient A_α of the potential/density expansion is a linear functional of f_1, and $A[f_-] = 0$, the derivative of f_+ in equation (2.24) can be replaced by a derivative of A_α to yield

$$j_\alpha[f_1](t) = -i\frac{\partial A_\alpha}{\partial t}.$$ (2.25)

2.4.1 Energy of a disturbance

When we use equation (2.8) to eliminate $|\mathbf{x} - \mathbf{x}'|$ from equation (5.130) of Binney and Tremaine (2008), we find that the energy associated with a linearised disturbance is

$$E[f_1] = \tfrac{1}{2}\left\{ \int \frac{\mathrm{d}^6\mathbf{w}}{|f_0'|} f_1^2 - \frac{\xi}{\mathcal{E}} \sum_\alpha \left| \int \mathrm{d}^6\mathbf{w}\, \Phi^{(\alpha)}(\mathbf{x}) f_1(\mathbf{w}) \right|^2 \right\}$$
$$= \tfrac{1}{2}\left\{ \langle f_-|f_-\rangle + \langle f_+|f_+\rangle - \xi\mathcal{E} \sum_\alpha \left| A_\alpha[f_+] \right|^2 \right\}.$$ (2.26)

In the first line of this equation, the first and second terms on the right quantify the potential and kinetic energies of the perturbation, respectively. Hence the inner product $\langle f_1|f_1\rangle$ gives twice a perturbation's kinetic energy. In the last term, A_α is independent of the degree of self-gravity ξ, so the final term is proportional to ξ as it should be.

In Appendix 2.9 we show that K is Hermitian and that

$$\langle f_-|K|f_-\rangle = |\omega^2|\left\{ \langle f_+|f_+\rangle - \xi\mathcal{E} \sum_\alpha \left| A_\alpha[f_+] \right|^2 \right\}.$$ (2.27)

When we use this equation to simplify equation (2.26), we discover the energy associated with an eigenfunction of K identically vanishes in the unstable case $\omega^2 < 0$, and in the stable case is

$$E[f_1] = \langle f_-|f_-\rangle = -(2\pi)^3 \int \frac{\mathrm{d}^3\mathbf{J}}{f_0'} \sum_\mathbf{k} |f_{\mathbf{k}-}|^2,$$ (2.28)

where $f_{\mathbf{k}-}$ denotes the component of f_- that fluctuates in angle-space as $\exp i\mathbf{k}\cdot\boldsymbol{\theta}$. Remarkably, the degree of self-gravity ξ does not appear in equation (2.28). Changes to ξ do affect E, however, by changing the $f_{\mathbf{k}-}$. The right side of equation (2.28) is inherently positive and vanishes only if f_- vanishes.

Moreover, by equation (2.14) $\omega f_+ = \mathbf{k}\cdot\mathbf{\Omega} f_-$, so f_+ must also vanish if f_- does. Hence *the energy of every mode of a stable ergodic system is positive.*

Since K is Hermitian it has a complete set of orthogonal eigenfunctions. Expressing an arbitrary disturbance as a linear combination of eigenfunctions $f_- = \sum_\beta c_\beta f_-^{(\beta)}$ and inserting this expansion into equation (2.28), we conclude that the disturbance's energy is the sum of the energies of its component eigenfunctions

$$E[f_1] = \sum_\beta |c_\beta|^2 \langle f_-^{(\beta)} | f_-^{(\beta)} \rangle. \tag{2.29}$$

In view of these results it is natural to identify the *true modes* of an ergodic system with the eigenfunctions of K. It follows that the frequencies of true modes are either real or pure imaginary. No true mode of an ergodic system has negative energy; oscillatory modes have positive energy and growing/decaying modes (if any) have zero energy. Equation (2.26), from which we started, does not make evident the non-negativity of energies.

2.4.2 Antonov's conserved quantity

The rather involved argument just given starts from equation (2.26) for the energy of a disturbance, which Binney and Tremaine (2008) derive by considering the work that must be done to initiate the disturbance. A simpler argument based on the Hermitian nature of K yields the closely related conserved quantity,

$$\widetilde{E} = \tfrac{1}{2}\Big(\langle \dot{f}_- | \dot{f}_- \rangle + \langle f_- | K | f_- \rangle \Big). \tag{2.30}$$

Indeed,

$$2\frac{\mathrm{d}\widetilde{E}}{\mathrm{d}t} = \langle \ddot{f}_- | \dot{f}_- \rangle + \langle \dot{f}_- | \ddot{f}_- \rangle + \langle \dot{f}_- | K | f_- \rangle + \langle f_- | K | \dot{f}_- \rangle = 0. \tag{2.31}$$

An eigenfunction of K with eigenvalue ω^2 has $\widetilde{E} = \omega^2 \langle f_- | f_- \rangle$, so in this case conservation of $\widetilde{E}$ implies conservation of $E = \widetilde{E}/\omega^2$. $\widetilde{E}$ does not have the dimensions of energy, however.

Since $\langle \dot{f}_- | \dot{f}_- \rangle > 0$, instability, and thus systematic growth of f_-, is excluded by conservation of $\widetilde{E}$ unless $\langle f_- | K | f_- \rangle < 0$ for some function $f_-(\mathbf{x}, \mathbf{v})$.

In fact positivity of $\langle f_- | K | f_- \rangle$ for all functions in the natural space is a necessary and sufficient condition for stability (Laval et al. 1965, Kulsrud and Mark 1970).

2.5 DFs of van Kampen modes

If the system is stable, the Hamiltonian's time-reversal invariance ensures that ω is real because the existence of exponentially decaying solutions would imply the existence of growing solutions. van Kampen (1955) applied these arguments to an electrostatic plasma and deduced some properties of the normal modes of a plasma, which are known as van Kampen modes. The corresponding modes of a stellar system have received little attention, although Vandervoort (2003) derived some of their properties. We now examine the van Kampen modes of a stellar system in their role as eigenfunctions of the operator K.

2.5.1 Operators that commute with K

Finding the eigenfunctions of K is facilitated by identifying operators that commute with K and seeking eigenfunctions of K that are also eigenfunctions of these operators. K is the operator associated with the time-translation invariance of the underlying equilibrium. The Hamiltonian H is invariant under increments in the angle variables θ_i but in the presence of self gravity ($\xi > 0$), K does not share the invariance with respect to increments in θ_1 and θ_2 conjugate to the radial action J_r and the modulus $L \equiv |\mathbf{L}|$ of the angular momentum vector $\mathbf{L}$ because $\Phi_1(\mathbf{x})$ lacks this invariance. Fortunately, K is always invariant under increments of the angle variable θ_3 conjugate to L_z. To see this it is best to return to the definition of K in equation (2.18). Since $[H, L_i] = 0$, we have

$$[Kf_-, L_z] = -[[[f_-, L_z], H], H] - f_0'(H)$$
$$\times \left[\int \mathrm{d}^6 \mathbf{w}' \left[\frac{\xi G}{|\mathbf{x}' - \mathbf{x}|}, L_z \right] [f_-(\mathbf{w}'), H(\mathbf{w}')], H(\mathbf{w}) \right]. \qquad (2.32)$$

The operator $[., L_z]$ rotates the orbital plane on which $\mathbf{x}$ lies (by incrementing the argument of the ascending node Ω). The operator $[., L_z + L_z']$, where L_z' operates on $\mathbf{x}'$, rotates $\mathbf{x}$ and $\mathbf{x}'$ through the same angle, so[3]

$$[|\mathbf{x}' - \mathbf{x}|, L_z + L_z'] = 0. \tag{2.33}$$

Hence taking advantage of the fact that for any f, g, h, $\int \mathrm{d}^6\mathbf{w} \, [f, g]h = \int \mathrm{d}^6\mathbf{w} \, f[g, h]$, we have

$$\begin{aligned}
\int \mathrm{d}^6\mathbf{w}' \, & \Big[\frac{\xi G}{|\mathbf{x}' - \mathbf{x}|}, L_z\Big][f_-(\mathbf{w}'), H(\mathbf{w}')] \\
&= -\int \mathrm{d}^6\mathbf{w}' \, \Big[\frac{\xi G}{|\mathbf{x}' - \mathbf{x}|}, L_z'\Big][f_-(\mathbf{w}'), H(\mathbf{w}')] \\
&= \int \mathrm{d}^6\mathbf{w}' \, \frac{\xi G}{|\mathbf{x}' - \mathbf{x}|}\Big[[f_-(\mathbf{w}'), H(\mathbf{w}')], L_z'\Big] \\
&= \int \mathrm{d}^6\mathbf{w}' \, \frac{\xi G}{|\mathbf{x}' - \mathbf{x}|}\Big[[f_-(\mathbf{w}'), L_z'], H(\mathbf{w}')\Big].
\end{aligned} \tag{2.34}$$

When this result is used in equation (2.32), we obtain

$$\big[Kf_-, L_z\big] = K[f_-, L_z], \tag{2.35}$$

so K commutes with the operator $[., L_z]$. From these commutations it follows that the eigenfunctions of K provide representations of the group of translations around tori that is generated by $[., L_z]$. This compact Abelian group has only one-dimensional irreps, which can be reduced to multiplication by $\mathrm{e}^{\mathrm{i}\alpha}$. Hence the eigenfunctions of A can be indexed by the integer n_3 associated with θ_3.

2.5.2 Derivation of the DF

Bearing in mind that for an eigenfunction equation (2.25) yields $j_\alpha[f_1] = -\omega A_\alpha$, from equation (2.23) we have for an eigenfunction

$$-\omega^2 f_- = \Big(\mathbf{\Omega} \cdot \frac{\partial}{\partial \boldsymbol{\theta}}\Big)^2 f_- - \mathrm{i}\xi\omega f_0'(H) \sum_\alpha A_\alpha \mathbf{\Omega} \cdot \frac{\partial}{\partial \boldsymbol{\theta}} \Phi^{(\alpha)}(\mathbf{x}). \tag{2.36}$$

---

[3]The operator $[., L]$ rotates the orbit within the orbital plane while holding constant θ_1, so in the Keplerian case it rotates the ellipse rather than moving the star along its ellipse. If $\mathbf{w}$ and $\mathbf{w}'$ lie on the same orbital plane, rotating both orbits within the plane will leave $|\mathbf{x}' - \mathbf{x}|$ invariant, but if the points lie on different planes, $|\mathbf{x}' - \mathbf{x}|$ will not be invariant. Hence $[|\mathbf{x}' - \mathbf{x}|, L] \neq 0$.

Now we apply the derivative $\mathbf{\Omega} \cdot \partial/\partial\boldsymbol{\theta}$ to both sides and use equations (2.14) and (2.19) to eliminate f_- in favour of f_+. Collecting terms with f_+ on the left we then have

$$\left\{\omega^2 + \left(\mathbf{\Omega} \cdot \frac{\partial}{\partial\boldsymbol{\theta}}\right)^2\right\} f_+ = \xi f_0'(H)\sum_\alpha A_\alpha \left(\mathbf{\Omega} \cdot \frac{\partial}{\partial\boldsymbol{\theta}}\right)^2 \Phi^{(\alpha)}(\mathbf{x}). \qquad (2.37)$$

Fourier decomposed in angle variables this becomes

$$\left\{\omega^2 - (\mathbf{n} \cdot \mathbf{\Omega})^2\right\} f_{\mathbf{n}+} = -\xi f_0'(H)(\mathbf{n} \cdot \mathbf{\Omega})^2 \sum_\alpha A_\alpha \Phi_{\mathbf{n}}^{(\alpha)}. \qquad (2.38)$$

This equation is analogous to the standard equation for the Laplace-transformed DF $\overline{f} \equiv \int_0^t \mathrm{d}t' \, f\mathrm{e}^{\mathrm{i}\omega t'}$ with $\Im(\omega) > 0$,

$$\mathrm{i}(\mathbf{n} \cdot \mathbf{\Omega} - \omega)\overline{f}_{\mathbf{n}}(\mathbf{J}, \omega) = -f_0'(H)\mathbf{n} \cdot \mathbf{\Omega} \sum_\alpha \overline{A}_\alpha(\omega)\Phi_{\mathbf{n}}^{(\alpha)} + f_{\mathbf{n}}(\mathbf{J}, 0) \qquad (2.39)$$

in that it relates the DF to the driving potential, but there are two significant differences:

1 The rhs of equation (2.38) does not contain an initial condition analogous to $\hat{f}_{\mathbf{n}}(\mathbf{J}, 0)$ on the rhs of (2.39). It is not there because we are seeking a normal mode rather than the solution to an initial-value problem.

2 Equation (2.38) starts with a factor $\omega^2 - (\mathbf{n} \cdot \mathbf{\Omega})^2$ while equation (2.39) starts with $\mathrm{i}(\mathbf{n} \cdot \mathbf{\Omega} - \omega)$.

Before we divide equation (2.38) by $\omega^2 - (\mathbf{n}\cdot\mathbf{\Omega})^2$, we must recognise that when it vanishes, which for a range of frequencies it will over 2d resonant surfaces in action space, $f_{\mathbf{n}+}$ is unconstrained. This being so, after we have divided by $\omega^2 - (\mathbf{n} \cdot \mathbf{\Omega})^2$, we should add a function that is non-zero only on resonant surfaces. Then we have

$$-f_{\mathbf{n}+}(\mathbf{J}) = \xi f_0'(H)\frac{(\mathbf{n} \cdot \mathbf{\Omega})^2}{\omega^2 - (\mathbf{n} \cdot \mathbf{\Omega})^2}\sum_\alpha A_\alpha \Phi_{\mathbf{n}}^{(\alpha)}(\mathbf{J})$$

$$+ g_{\mathbf{n}}(\mathbf{J})\delta(\omega^2 - (\mathbf{n} \cdot \mathbf{\Omega})^2), \qquad (2.40)$$

where $g_{\mathbf{n}}$ is an arbitrary function. Adding this term makes it possible for $f_{\mathbf{n}+}$ to take whatever value $g_{\mathbf{n}}$ specifies on the resonant surfaces. (The values

taken by $g_\mathbf{n}$ off resonant surfaces are immaterial.) Multiplying equation (2.40) by $\int \mathrm{d}^6\mathbf{w}\, \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}}\Phi^{(\alpha')*}(\mathbf{x})$ and summing over $\mathbf{n}$, we get

$$
\begin{aligned}
\mathcal{E}A_{\alpha'} &= \int \mathrm{d}^3\mathbf{J}\mathrm{d}^3\boldsymbol{\theta} \sum_\mathbf{n} \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}}\Phi^{(\alpha')*}(\mathbf{x}) \\
&\quad \times \left\{ \xi f_0'(H)\frac{(\mathbf{n}\cdot\boldsymbol{\Omega})^2 \sum_\alpha A_\alpha \Phi_\mathbf{n}^{(\alpha)}}{\omega^2 - (\mathbf{n}\cdot\boldsymbol{\Omega})^2} + g_\mathbf{n}(\mathbf{J})\delta(\omega^2 - (\mathbf{n}\cdot\boldsymbol{\Omega})^2) \right\} \\
&= (2\pi)^3 \mathcal{P}\!\int \mathrm{d}^3\mathbf{J}\sum_\mathbf{n} \left\{ \xi f_0' \frac{(\mathbf{n}\cdot\boldsymbol{\Omega})^2}{\omega^2 - (\mathbf{n}\cdot\boldsymbol{\Omega})^2}\sum_\alpha A_\alpha \Phi_\mathbf{n}^{(\alpha')*}\Phi_\mathbf{n}^{(\alpha)} \right. \\
&\quad \left. + \Phi_\mathbf{n}^{(\alpha')*}g_\mathbf{n}(\mathbf{J})\delta(\omega^2 - (\mathbf{n}\cdot\boldsymbol{\Omega})^2) \right\},
\end{aligned}
\tag{2.41}
$$

where the integral over $\mathbf{J}$ is a principal value in the sense that actions at which $\mathbf{n}\cdot\boldsymbol{\Omega} = \pm\omega$ are to be excluded and the large values of the integrand as such points are approached largely cancel during integration. Equation (2.41) has the form

$$
\sum_\alpha \mathcal{M}_{\alpha'\alpha}A_\alpha = -B_{\alpha'},
\tag{2.42}
$$

where[4]

$$
\begin{aligned}
\mathcal{M}_{\alpha'\alpha}(\omega) &\equiv \delta_{\alpha'\alpha} \\
&\quad - \frac{(2\pi)^3\xi}{\mathcal{E}}\mathcal{P}\!\int \mathrm{d}^3\mathbf{J}\, f_0'(H) \sum_\mathbf{n} \frac{(\mathbf{n}\cdot\boldsymbol{\Omega})^2}{\omega^2 - (\mathbf{n}\cdot\boldsymbol{\Omega})^2}\Phi_\mathbf{n}^{(\alpha')*}\Phi_\mathbf{n}^{(\alpha)} \\
B_{\alpha'} &\equiv -\frac{(2\pi)^3}{\mathcal{E}}\int \mathrm{d}^3\mathbf{J}\sum_\mathbf{n} \Phi_\mathbf{n}^{(\alpha')*}(\mathbf{J})g_\mathbf{n}(\mathbf{J})\delta(\omega^2 - (\mathbf{n}\cdot\boldsymbol{\Omega})^2).
\end{aligned}
\tag{2.43}
$$

We will see below that in the case of a stable system, $\mathcal{M}(\omega)$ has an inverse for any real ω. Consequently, for any real ω and B_α there is always a unique corresponding $A_\alpha(\omega)$. Hence given ω and $\mathbf{n}$, we can determine the A_α for any

---

[4]Our matrix $\mathcal{M}$ is analogous to the matrix $\boldsymbol{\epsilon}$ of Hamilton et al. (2018) rather than their $\mathbf{M} = \mathbf{I} - \boldsymbol{\epsilon}$. The integrand in our $\mathcal{M}$ differs from that of Hamilton et al. (2018) in that frequencies occur squared because we are working with an operator that is second- rather than first-order in time. The polarisation and response operators defined in Chapter 5 of Binney and Tremaine (2008) also involve frequencies rather than their squares. Similar operators can be derived from our $\mathcal{M}$ by decomposition of our integral by partial fractions.

function $g_{\mathbf{n}}(\mathbf{J})$ on the surface $\mathbf{n} \cdot \mathbf{\Omega} = \omega$. By equation (2.6), these coefficients describe the spatial structure that g generates:

$$\Phi_A(\mathbf{x}) = \xi \sum_\alpha A_\alpha \Phi^{(\alpha)}(\mathbf{x}) \quad ; \quad \rho_A(\mathbf{x}) = \sum_\alpha A_\alpha \rho^{(\alpha)}(\mathbf{x}). \tag{2.44}$$

The B_α turn out to be the analogous expansion coefficients of the density

$$\rho_B \equiv \int \mathrm{d}^3\mathbf{v} \sum_{\mathbf{n}} g_{\mathbf{n}}(\mathbf{J}) \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}} \delta(\omega^2 - (\mathbf{n} \cdot \mathbf{\Omega})^2). \tag{2.45}$$

Indeed,

$$-\frac{1}{\mathcal{E}} \int \mathrm{d}^3\mathbf{x} \, \Phi^{(\alpha)*}(\mathbf{x}) \rho_B = -\frac{1}{\mathcal{E}} \sum_{\mathbf{n}\mathbf{n}'} \int \mathrm{d}^3\mathbf{J} \int \mathrm{d}^3\boldsymbol{\theta} \, \Phi^{(\alpha)*}_{\mathbf{n}'}(\mathbf{J}) \mathrm{e}^{-\mathrm{i}\mathbf{n}'\cdot\boldsymbol{\theta}}$$
$$\times g_{\mathbf{n}}(\mathbf{J}) \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}} \delta(\omega^2 - (\mathbf{n} \cdot \mathbf{\Omega})^2)$$
$$= B_\alpha. \tag{2.46}$$

We shall call modes with real frequencies and non-zero $g_{\mathbf{n}}$ *van Kampen* modes.

At $\omega^2 < 0$, $B_\alpha = 0$ because the argument of the δ-function in the definition of B_α cannot vanish. So given $\omega^2 < 0$ an associated DF can be found only if $|\mathcal{M}| = 0$. Thus there may be isolated pure imaginary frequencies $\pm \mathrm{i}\omega_0$ at which a perturbation can grow/decay exponentially. We call such modes *classical modes*. They are distinct from damped Landau 'modes', which have frequencies below the real axis and not on the imaginary axis.

Note that to prove the system's stability it suffices to show that $|\mathcal{M}|$ has no zeroes on the imaginary axis. In any case, the normal-mode frequencies are confined to a continuum of real values (van Kampen modes) with the possible addition of discrete pure imaginary values (classical modes).

2.5.3 Relation to Landau modes

In a conventional normal-mode analysis we derive an equation $M\mathbf{a} = 0$ which is homogeneous in the disturbance's amplitude $\mathbf{a}$ with the consequence that non-trivial solutions exist only when the determinant of the matrix M vanishes. The dispersion relation is the condition for $|M(\omega)|$ to vanish. Our recognition that there can be non-trivial distributions of stars on resonant tori gives rise to

non-vanishing B_α, and thus causes A_α to satisfy an inhomogeneous equation analogous to $M\mathbf{a} = \mathbf{b}$ that can be satisfied whenever $B_\alpha \neq 0$, that is, at any frequency $\omega = \mathbf{n} \cdot \boldsymbol{\Omega}(\mathbf{J})$ for some $\mathbf{n}$ and $\mathbf{J}$. Hence there is no dispersion relation associated with the true modes of a stellar system. In general there will be complex frequencies ω_0 at which $|\mathcal{M}(\omega_0)| = 0$, but unless ω_0 lies on an axis of the complex plane, it cannot be the frequency of a true mode because all true modes have real ω^2.

Landau modes occur at the frequencies ω_0 at which a matrix $M(\omega)$ has vanishing determinant. The Landau matrix M is closely related to $\mathcal{M}$, and if $|M(\omega_0) = 0$ then $|\mathcal{M}(\pm\omega_0)| = 0$ also. Thus every Landau mode is associated with the possibility of solving equation (2.42) with $B_\alpha = 0$. Yet when ω_0 lies on neither axis of the complex plane, such a solution should not be included in the set of true modes for two reasons:

i) The solution cannot be an eigenfunction of K because ω_0^2 is not real, so it falls outside the complete set formed by the true modes. (It follows that it can be written as a sum of the true modes.)

ii) The solution associated with one of $\pm\omega_0$ will grow exponentially, so every system would be unstable if these solutions were included in the complete set of true modes.

If a system is stable, the frequencies of its Landau modes all lie below the real axis. In Section 2.5.2 we used this fact to argue that given any real ω, equation (2.42) has a unique solution for A_α given B_α. Time-reversal symmetry is responsible for $|\mathcal{M}|$ vanishing above the real axis whenever it vanishes below the real axis and the failure of the determinant of the Landau matrix M to behave in the same way is a consequence of the the violation of time-reversal symmetry inherent in the initial-value problem that leads to M.

2.5.4   Modes and dressing

To understand the physical reality that underlies this mathematics, consider that in the absence of self gravity ($\xi = 0$), a non-trivial distribution of stars

with respect to $\boldsymbol{\theta}$ on resonant surfaces will generate oscillations in the density at frequency ω that will persist for ever. The δ-function component of the DF (2.40) of a van Kampen mode represents this phenomenon, and the B_α quantify the spatial form of this driving structure. When $\xi > 0$, these oscillations affect the dynamics of all stars, including non-resonant stars. The regular part of the mode's DF (2.40) describes these sympathetic oscillations of non-resonant stars. The A_α quantify the spatial structure of this "dressed" response to the driver g.

The functions $g_{\mathbf{n}}$ are arbitrary, so we may consider the case in which $g_{\mathbf{n}}$ vanishes for all vectors but one, $\mathbf{N}$. Consider now the effect of multiplying equation (2.40) by $\int \mathrm{d}^6\mathbf{w}\, \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}}\Phi^{(\alpha')*}(\mathbf{x})$ as in the derivation of equation (2.41) but now *not* summing over $\mathbf{n}$. Then we have

$$C_{\alpha'\mathbf{n}} = (2\pi)^3 \mathcal{P}\int \mathrm{d}^3\mathbf{J}\,\Phi_{\mathbf{n}}^{(\alpha')*}\left\{\xi f_0' \frac{(\mathbf{n}\cdot\boldsymbol{\Omega})^2}{\omega^2 - (\mathbf{n}\cdot\boldsymbol{\Omega})^2}\sum_\alpha A_\alpha \Phi_{\mathbf{n}}^{(\alpha)}\right.$$
$$\left. + g_{\mathbf{n}}(\mathbf{J})\delta(\omega^2 - (\mathbf{n}\cdot\boldsymbol{\Omega})^2)\right\}, \tag{2.47}$$

where

$$C_{\alpha'\mathbf{n}} \equiv -\int \mathrm{d}^6\mathbf{w}\,\Phi^{(\alpha)*}(\mathbf{x})\mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}}f_{\mathbf{n}+}, \tag{2.48}$$

so

$$A_\alpha = \frac{1}{\mathcal{E}}\sum_{\mathbf{n}} C_{\alpha\mathbf{n}} \tag{2.49}$$

When $\mathbf{n} \neq \mathbf{N}$, the right side of equation (2.47) has only the term proportional to ξ, so would vanish with ξ. That is, without self-gravity, $C_{\alpha\mathbf{n}} \neq 0$ only for $\mathbf{n} = \mathbf{N}$; in this case the van Kampen modes could be labelled by $\mathbf{N}$. In the presence of self-gravity, we have no reason to expect $C_{\alpha\mathbf{n}}$ to vanish for $\mathbf{n} \neq \mathbf{N}$ because equation (2.49) includes a contribution to A_α from $C_{\alpha\mathbf{N}} \neq 0$. Self-gravity has this impact because it stops K commuting with the operators $[., J_r]$ and $[., L]$ as discussed above. From the fact that K does commute with the operator $[., L_z]$ it follows that even when $\xi > 0$, $A_{\alpha\mathbf{n}} = 0$ unless $n_3 = N_3$.

When a Landau mode is weakly damped, $|\mathcal{M}| = 0$ at a frequency ω_0 that lies just below the real axis, and in consequence $|\mathcal{M}|$ is small on the real axis just above this zero. In view of equation (2.42), $|\mathbf{A}|/|\mathbf{B}|$ will be large in these circumstances. That is, van Kampen modes with frequencies close to $\Re(\omega_0)$ are "heavily dressed". This is the true significance of Landau modes. Put differently, while van Kampen modes exist at any ω, they have a bigger footprint in action space at frequencies that lie close to those of Landau modes.

Whereas a gas ball has at most a finite number of normal modes at a countable number of frequencies, a cluster has an infinite number of normal modes at every frequency. This difference is a consequence of the likely completeness of normal modes in each system (Section 2.6) and the fact that much more information is required to specify the DF of a cluster than the state of a gas ball: the disturbance has adiabatically deformed the latter from its equilibrium, so the velocity distribution remains Maxwellian and one only has to specify a mean and dispersion at each location $\mathbf{x}$. In a cluster we need to specify the DF in six-dimensional phase space.

Whereas the Hermiticity of K ensures that van Kampen modes for different ω are orthogonal, it falls to us to select from all modes for any given ω a complete set of mutually orthogonal modes. That is, to select a set of functions $g(\mathbf{J})$ that generate modes that are orthogonal in the sense $\langle g_i | g_j \rangle = \delta_{ij}$. Since K commutes with $[., L_z]$, we can require modes to be eigenfunctions of this Hermitian operator, and identification of a complete set of modes is reduced to finding an orthogonal set of vectors whose components are indexed by n_1 and n_2. Rather than solving this problem, in the forthcoming third contribution we show how an arbitrary state of a stellar system can be decomposed into its constituent modes. That is, to express a given perturbation at time zero, $F(\mathbf{w}, 0)$, in the form

$$F(\mathbf{w}, 0) = \int d\omega \, f(\mathbf{w}, \omega), \tag{2.50}$$

where $f(\mathbf{w}, \omega)$ is a van Kampen mode with frequency ω. This done, the state

of the perturbation at any other time can be obtained as

$$F(\mathbf{w}, t) = \int \mathrm{d}\omega \, f(\mathbf{w}, \omega) \, \mathrm{e}^{-\mathrm{i}\omega t}. \tag{2.51}$$

These integrals over ω, eliminate the principal-value and Dirac δ-function symbols in equation (2.40) for f.

2.5.5 Energy of a mode

Equation (2.28) says that the energy of a mode is just the norm of the odd part of its DF, and equation (2.40) gives the even part of a mode's DF from which the odd part follows trivially. The next step is to substitute from the second of these equations into the first and express the mode's energy in terms of its parameters $g_{\mathbf{n}}$ and its potential $\Phi[f]$. This exercise proves long, and is made intricate by the singular denominator $\omega^2 - (\mathbf{n} \cdot \boldsymbol{\Omega})^2$ in equation (2.40) for the DF.

In Appendix 2.10 we compute $\langle f | \widetilde{f} \rangle$ for modes with frequencies ω and $\widetilde{\omega}$. The result is[5]

$$\langle f_- | \widetilde{f_-} \rangle = -(2\pi)^3 \, \delta(\omega^2 - \widetilde{\omega}^2) \int \mathrm{d}^3 \mathbf{J} \sum_{n_1, n_2}$$
$$\left(\pi^2 \omega^4 f_0' \Phi_{\mathbf{n}}^*[f] \Phi_{\mathbf{n}}[\widetilde{f}] + \frac{g_{\mathbf{n}}^* \widetilde{g}_{\mathbf{n}}}{f_0'} \right) \delta((\mathbf{n} \cdot \boldsymbol{\Omega})^2 - \omega^2). \tag{2.52}$$

A feature of this expression is that it has no reference to potential/density basis functions. The factor $\delta(\omega^2 - \widetilde{\omega}^2)$ on the rhs reflects the orthogonality of modes of different frequencies that follows from the Hermiticity of K.[6] A remarkable feature of equations (2.52) is that its action-space integral is confined to the resonant tori, even though the modes very much involve non-resonant stars.

When we set $\widetilde{f} = f$, the inner product is divergent for finite g because finite g generates an f_+ that diverges as the resonant surface in action space is approached, and the divergences on opposite sides of the surface do not cancel because energy density is proportional to f^2. This result signals that we can

---

[5]In equation (2.52) n_3 has the same, fixed value for both f and $\widetilde{f}$.
[6]The appearance of ω^2 rather than ω reflects time-reversal symmetry. From a practical perspective, it ensures that real time dependence ($\cos \omega t$) is possible.

have only an infinitesimal number of stars on any resonant surface. When we use equation (2.52) to compute the energy of a physical disturbance (2.50), we find

$$
\begin{aligned}
E[F] = \langle F_- | F_- \rangle &= \left\langle \int \mathrm{d}\omega\, f_-(\omega) \,\middle|\, \int \mathrm{d}\omega'\, f_-(\omega') \right\rangle \\
&= \int \mathrm{d}\omega\, \mathrm{d}\omega'\, \langle f_-(\omega) | f_-(\omega') \rangle \\
&= -(2\pi)^3 \sum_{\mathbf{n}} \int \frac{\mathrm{d}\omega}{2\omega} \int \mathrm{d}^3\mathbf{J} \left(\pi^2 \omega^4 f_0' |\Phi_{\mathbf{n}}[f]|^2 + \frac{|g_{\mathbf{n}}|^2}{f_0'} \right) \\
&\qquad\qquad\qquad\qquad \times \delta((\mathbf{n} \cdot \mathbf{\Omega})^2 - \omega^2) \\
&= -\frac{(2\pi)^3}{4} \sum_{\mathbf{n}} \int \mathrm{d}^3\mathbf{J} \left(\pi^2 (\mathbf{n} \cdot \mathbf{\Omega})^2 f_0' |\Phi_{\mathbf{n}}[f]|^2 + \frac{|g_{\mathbf{n}}|^2}{(\mathbf{n} \cdot \mathbf{\Omega})^2 f_0'} \right).
\end{aligned}
$$

$$(2.53)$$

There is a striking similarity between the first term in the expression (2.53) for E and the expression $\rho_E = \frac{1}{2}\omega^2 \epsilon_0 A^2$ for the energy density contributed by an electromagnetic wave with vector-potential amplitude A. In the electromagnetic case one factor of ω arises from the quantisation condition $E = \hbar\omega$ and the other arises from the canonical momentum $\omega\mathbf{A}$ of the field $\mathbf{A}$. The second term in equation (2.53) has a different structure, however, and this term is arguably more important than the first because $\Phi[f]$ is driven by g.

A question to ask is why the coefficient of $\left|\Phi_{\mathbf{n}}[f]\right|^2$ in equation (2.53) is positive, given that gravitational potential energy is inherently negative. The explanation must be that this term encapsulates all the energy, kinetic as well as potential, that is tied up in the disturbance in non-resonant stars that is excited by the resonant stars. In the absence of self-gravity, the non-resonant stars are not disturbed, so this contribution to the energy vanishes with $\Phi[f]$.

The second term in the integrand of equation (2.53) is ultimately limited by Poisson noise and can be considered a given, while the first term depends on the system's dynamics. The ratio of the two terms is proportional to $(\mathbf{n} \cdot \mathbf{\Omega})^4 f_0'^2$, so the relative contributions to E from the resonant driver g and the non-

resonant response Φ are sensitive to this factor. In principle $|\mathbf{n} \cdot \boldsymbol{\Omega}|$ can be made as small as we please at given $\mathbf{J}$, but only by going to large $|\mathbf{n}|$, and at large $|\mathbf{n}|$ the projection of $g_{\mathbf{n}}(\mathbf{J})$ into real space (measured by B_α) will be small and thus the response (measured by A_α and $\Phi[f]$) will be small. Hence the first term in the integrand of equation (2.53) will be significant only for small $|\mathbf{n}|$. The fundamental dipole mode satisfies this criterion.

This consideration draws attention to short vectors $\mathbf{n}$ that make $|\mathbf{n} \cdot \boldsymbol{\Omega}|$ small (if it is small on any torus, it will be small on many tori) because for these vectors the noise component $g_{\mathbf{n}}$ generates a response at the least cost in energy, so the response is likely to be large. We saw above, moreover, that the response is enhanced when $\mathbf{n} \cdot \boldsymbol{\Omega}$ is close to the real part of the frequency of a weakly damped Landau mode, because then $|\mathcal{M}|$ is small. The fundamental dipole mode has been shown to be weakly damped in typical models (Weinberg 1994, Saha 1991, Hamilton et al. 2018).

2.5.6 Initialisation of N-body models

Suppose we set up a cluster by randomly sampling an analytic DF. When the selection is complete, the actual DF will differ from the analytic DF by virtue of Poisson noise, so the $g_{\mathbf{n}}$ will be non-zero and the cluster's van Kampen modes will be excited. The coefficients B_α that quantify the noisiness of the density distribution are unambiguously fixed by the Monte-Carlo selection, but the potential that is generated from them $\xi \sum_{\alpha\beta} \Phi^{(\alpha)} \mathcal{M}^{-1}(\xi, \omega) B_\beta$ will vary with the degree of self-gravity ξ. Hence the DFs of the modes that sum to the sampled phase-space distribution will depend on ξ, but their sum must produce the DF sampled regardless of ξ. When $\xi = 1$, the modes are heavily dressed and yet produce the same small (Poisson) fluctuations in density as in the case $\xi = 0$ of vanishing self-gravity because the contributions of different modes cancel to a considerable extent. This cancellation is particularly pronounced in the case of low-order modes (small $|\mathbf{n}|$), and it occurs because when $\xi = 1$ the phases of modes are correlated, whereas when $\xi = 0$ they are probably uncorrelated.

Once we start moving the stars with ξ set to unity, the phase differences between modes with different frequencies will tend towards uniform distribution in $(0, 2\pi)$ and cancellations between perturbations to the density will diminish. Consequently, the heavily dressed individual modes will become manifest and the system will become less spherical as Lau & Binney (2019) found empirically. By contrast, when stars are moved in the analytic potential, the initially uniformly distributed phases of the modes obtained with $\xi = 0$ remain uniformly distributed and no significant change in the density fluctuations will be observed.

The larger the value of ξ, the larger will be the values of A_α that are generated by the given B_α, so the greater will be the departures from spherical symmetry once the phases of modes have decorrelated.

To obtain a self-consistent realisation of a self-gravitating system one needs to excite the modes for $\xi = 1$ with random phases, and it is not clear how this can be done without computing the system's modes.

2.6   Completeness of modes

We have defined the true modes of a stellar system to be the eigenfunctions of Antonov's Hermitian operator K. In quantum mechanics it is conventional to assume that the eigenfunctions of any Hermitian operator form a complete set although proof of completeness requires the operator to be bounded (e.g. Dieudonné 1969, §11.5), which some operators of physical interest are not. Similarly, much of condensed-matter physics relies on Bloch's theorem that there is a complete set of stationary states for an electron in a crystal that have wavefunctions of the form $\psi(\mathbf{x}) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}u(\mathbf{x})$ with $u(\mathbf{x} + \mathbf{a}) = u(\mathbf{x})$ for any lattice vector $\mathbf{a}$. The standard derivation of Bloch's theorem (e.g. Elliott and G. 1989, §14.4) starts from the observation that if $\psi(\mathbf{x})$ is a stationary state, then so is $\psi(\mathbf{x} + \mathbf{a})$. Hence the stationary states of a given energy provide a representation of the Abelian translation group. Such groups only have one-dimensional irreducible representations, so the action of the group

can be reduced to multiplication by $e^{i\mathbf{k}\cdot\mathbf{a}}$. That is, any functions providing a representation of the translation group can be reduced to functions satisfying $\psi(\mathbf{x} + \mathbf{a}) = e^{i\mathbf{k}\cdot\mathbf{a}}\psi(\mathbf{x})$, a relation that is clearly satisfied by $\psi(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}}u(\mathbf{x})$.

Analogously, we might argue that the time-translation invariance of $\partial_t^2 + K$ implies that if $f(\mathbf{w}, t)$ satisfies $(\partial_t^2 + K)f = 0$, then so does $f(\mathbf{w}, t + \tau)$ and it follows that any set of solutions can be reduced to ones that satisfy $f(\mathbf{w}, t + \tau) = e^{-i\omega\tau}f(\mathbf{w}, t)$. Then setting t to zero we infer that solutions of the form $f(\mathbf{w}, \tau) = e^{-i\omega\tau}f(\mathbf{w})$ are complete.

The above arguments for the completeness of Bloch waves and eigenfunctions of A are open to the objection that the theorem regarding the decomposition of representations into irreducible representations requires the group to be compact, which translation groups are not.[7] Hence, the completeness of van Kampen modes cannot be rigorously established by the group-theoretic argument, although similar arguments are widely accepted in physics.

Case (1959) established the completeness of the van Kampen modes of an electrostatic plasma by direct demonstration that any DF $f(\mathbf{w})$ can be written as a sum of van Kampen modes. The corresponding exercise for stellar systems will be presented in the third contribution of this series.

2.7 Discussion

2.7.1 Particle dressing

The concept of particle dressing has been central to high-energy physics for over half a century, but has been unadopted in stellar dynamics until recently. In galactic dynamics it can be traced at least as far back as Julian and Toomre (1966), who showed that a gas cloud in a galactic disc would attract an entourage of passing stars $\sim$ten times more massive than itself. Toomre and Kalnajs (1991) showed that individual disc stars also enhance their masses ten-

---

[7]The reduction theorem applies only to unitary representations, which associates every group member g with a unitary operator T_g on a vector space. If a group is compact, Maschke's operator $S^2 = \sum_g T_g^\dagger T_g$ can be used to establish that any representation is isomorphic to a unitary representation. In the non-compact case the sum over g is ill-defined.

fold by attracting an (ever-changing) entourage of other stars. Sellwood and Carlberg (2014) showed that large-amplitude spiral structure emerges from Poisson noise through successive spiral instabilities, but this important process was only firmly connected to particle dressing by Fouvry et al. (2015). Hamilton (2021) made the connection between the BL equation and particle dressing clear via Rostoker's principle: that it is permissible to compute the effects of discreteness as from the interaction of uncorrelated but dressed particles (Rostoker 1964). Here we have interpreted a van Kampen mode as the result of dressing not one star but an ensemble of stars on a group of resonant tori.

Given that the CBE is the first equation in the BBGKY hierarchy of equations with the two-particle correlation function set to zero, the importance of dressing for the structure of van Kampen modes may seem paradoxical. The CBE is a mean-field approximation akin to the Weiss theory of magnetism, and embraces correlations that are induced by perturbing fields. Hence it embraces the dressing of resonant tori involved in van Kampen modes. Moreover, when a simulation is started, its DF is inevitably perturbed from the underlying analytic DF and thus its van Kampen modes are excited.

2.7.2 van Kampen vs Landau modes

Normal modes are perhaps the most ubiquitous tools in theoretical physics – quantum field theory has even taught us to see particles as excitations of normal modes of the vacuum. Modal analyses have played a significant role in stellar dynamics since the seminal work of Kalnajs (1965), Toomre (Toomre 1964; 1981) and later the prescient work of Weinberg (Weinberg 1993; 1994; 1998; 2001), but in all these studies the modes considered were those of Landau. These 'modes' lack key properties of true modes: (i) completeness in the sense that any initial condition can be expressed as a linear combination of modes, and (ii) additivity of energies. These two properties are essential for the use of modes in physics and engineering outside stellar dynamics. The contents pages of the two volumes of Fridman and Polyachenko (1984) explain the focus on

Landau modes: the community wanted to establish which equilibrium models are stable, rather than to investigate, as we do, the excursions that stable systems make around equilibrium.

Doubt is sometimes cast on the physical standing of van Kampen modes because their DFs contain a δ-function. Actually this feature is a natural consequence of their forming a continuum. Testable predictions of van Kampen modes will always emerge after integration over ω, just as in the familiar quantum-mechanical treatment of radiative transitions sensible results emerge only after integration over the frequency of the electromagnetic field,[8] and the δ-functions will disappear in the process. Landau modes are superpositions of van Kampen modes, and they decay as their constituent van Kampen modes drift apart in phase (Case 1959). If we run a decayed Landau mode back in time, the phases move back into alignment for a finite time before drifting apart, so the disturbance grows for only a limited time.

Hamilton and Heinemann (2020) take a fresh approach to relaxation in stellar systems that involves Landau modes in an essential way. Binney and Lacey (1988) showed that the diffusion tensor of the action-space Fokker-Planck equation follows immediately from the temporal power spectrum of the gravitational potential. Hamilton and Heinemann (2020) argue that both dressed two-particle interactions and normal modes of the entire system contribute to the power spectrum. They assume that the modes in question are Landau modes, which they imagine to reach an equilibrium level of excitation through their native damping being offset by constant excitation by Poisson noise. This picture involves a transfer of energy from Landau modes to the underlying heat bath, and then back to the Landau modes via Poisson noise. The mechanism by which Poisson noise draws energy from the heat bath is unclear.

In a simpler picture each van Kampen mode has a fixed amplitude and energy and a phase that advances at its own steady rate. Modes with fre-

---

[8]When deriving Fermi's golden rule, one analogously integrates over the energies of final states.

quencies that lie close to the real parts of a weakly damped Landau mode have large amplitudes because they are heavily dressed. From time to time their phases yield constructive interference and the Landau mode appears to be highly excited. The excitation decays as shifts in relative phase spoil the constructive interference. At a later time the phases again align favourably, and the process repeats.

2.7.3 Thermodynamics of star clusters

When a stellar system is born, its van Kampen modes are assigned particular amplitudes and phases. At birth the phases may be highly correlated, but they will decorrelate on a dynamical timescale. This decorrelation may manifest itself through the emergence of system-scale fluctuations in the density, wandering of the point of highest central density, (e.g. Lau and Binney 2019, Heggie et al. 2020). On a longer timescale non-linear terms in the CBE will mediate exchanges of energy between modes. We know that the amplitudes of modes are invariant at linear order, and we expect them to evolve at quadratic order in the perturbations. The two-body timescale is precisely the timescale associated with terms of quadratic oder (Chavanis 2012), so van Kampen modes are expected to exchange energy on the two-body timescale.

There is no reason to believe that when a cluster is first realised the amplitudes of its van Kampen modes conform to the Gibbs distribution. We expect exchanges of energy between modes to drive the distribution towards the Gibbs distribution and equipartition of energy between modes, which is to say that the actual DF is $F = \int \mathrm{d}\omega \, f(\omega)$ with $f(\omega)$ of van-Kampen form (2.40) and (cf. eqn 2.53)

$$E[f] = \langle f_- | f_- \rangle = \text{constant.} \tag{2.54}$$

Formally, modes exchange energy and equipartition can be approached on the same (two-body) timescale on which core collapse and evaporation change the mean-field model, but since core collapse occurs after ~ 300 central two-body times (Binney and Tremaine 2008, §7.5.3), it is plausible that a good

approximation to the Gibbs distribution will be achieved before core collapse occurs. A programme of work consistent with with this separation of timescales is to assume equipartition of energy and random phases between the modes of a particular mean-field model and to compare the observables predicted thus with observational data and N-body simulations.

In this regard it is instructive to compare the applicability of thermodynamics to star clusters and to classical systems that also have access to states of very high entropy, for example a mixture of two parts hydrogen and one part oxygen, or a diamond, both of which have higher entropy states (water vapour and graphite) that can only be reached by climbing over a significant energy barrier. On account of this barrier, a hydrogen/oxygen mixture and a diamond will extensively explore the configurations accessible with thermal energy regardless of the existence of states of much higher entropy. A stellar system is not denied access to states of high energy by an energy barrier but by low rates of energy transport.

Nevertheless, computing the thermodynamics of a cluster is feasible because the positivity of model energy ensures that the cluster's constant-energy surfaces have finite volume Ω, and it would be very interesting to examine its predictions.

In such a theory the DF would itself be a random variable in addition to the coordinates of stars, which are the random variables whose probability distribution the DF specifies. Testable predictions would emerge from the theory as double expectations: first $\langle \mathcal{O} \rangle_f = \int \mathrm{d}^6 \mathbf{w} \, f(\mathbf{w}) \mathcal{O}(\mathbf{w})$ and then an average of these averages weighted by the probability of each DF f. Hence a prerequisite of the theory is the ability to assign probabilities to DFs in a rational way. In particular, the probability assigned to a group of DFs must remain unchanged as a cluster evolves under the CBE. In standard statistical mechanics the analogous requirement, that a probability density on phase space be invariant under Hamiltonian evolution, is satisfied by making a priori probability proportional to the measure of phase-space volume $\mathrm{d}^N \mathbf{q} \, \mathrm{d}^N \mathbf{p}$ that

canonical coordinates $(\mathbf{q}, \mathbf{p})$ deliver. In Chapter 4 we extend this idea to the space of distribution functions by defining canonical coordinates for this space. It turns out that the energy of a van Kampen mode then takes the form of a sum of Hamiltonians of simple-harmonic oscillators.

2.7.4 Prior work

To our knowledge the van Kampen modes of a stellar system have previously been considered only by Vandervoort (2003), who followed van Kampen (1955) in deriving the modes directly from the CBE. Antonov (1961) obtained the second-order, Hermitian differential operator K by splitting the DF into parts even and odd in $\mathbf{v}$, but he was focused on proving his stability principle and did not show that the eigenfunctions of K are the van Kampen modes. He did not take advantage of angle-action variables or compute excitation energies. Polyachenko et al. (2021) discussed the relation of Landau and van Kampen modes in the context of the periodic cube, though principally in the unphysical case that the cube's mass exceeds the Jeans mass so the system is unstable. Their work makes very clear that Landau 'modes' lack essential properties of modes, and also illustrates how calculations in the complex plane obfuscates attempts at their physical interpretation: physically ill-motivated changes in the contour of integration over velocity give rise to solutions with radically different properties. In particular they show that Landau's choice of contour breaks the time-reversal symmetry of the underlying problem.

Polyachenko et al. (2021) introduced the nomenclature 'true mode' for a member of the complete set of modes and like us restricted the term 'van Kampen mode' to true modes with real frequencies. The principal point made by Polyachenko et al. (2021) is that the unstable Jeans mode at $\omega = \mathrm{i}y$ is accompanied by a decaying mode at $\omega = -\mathrm{i}y$. This is a trivial consequence of time-reversibility but Polyachenko et al. (2021) show that some effort is required to explain why Landau's analysis misses this mode.

2.8   Conclusions

As the completeness and precision of astronomical data grow (i.e. we approach being able to observe every star which lies within our observational range), the oscillations of stellar systems around equilibrium configurations will increase in observational significance. The way to produce theoretical predictions of these phenomena is to adapt the techniques of statistical mechanics to stellar dynamics. This chapter takes a step in this direction for ergodic stellar systems by focusing attention on the van Kampen modes of stellar systems, which have hitherto been eclipsed by Landau modes.

We showed for the first time that van Kampen modes of an ergodic system are the eigenfunctions with positive eigenvalues of Antonov's second-order Hermitian operator on phase space. In consequence, the true modes of an ergodic stellar system are either purely sinusoidal or exponentially growing/decaying; there are no over-stable modes or modes comprising decaying oscillations. The frequencies of oscillating modes form a continuum, and the DFs of these modes contain δ-functions which disappear when testable predictions are extracted by integrating over frequencies. Any exponentially growing/decaying modes are isolated in frequency space and their DFs do not contain δ-functions. The energy of an oscillating mode is just the norm of the odd part of its DF. From this it follows that these modes have positive energy. The energy of a growing/decaying mode is identically zero.

We interpreted van Kampen modes as dressed sets of resonant tori. How heavily they are dressed increases with the extent to which the system is self-gravitating, and with proximity in frequency space to a zero of the response matrix $\mathcal{M}(\omega)$ – the zeroes of this matrix come in pairs with each pair that is not on the imaginary axis associated with a Landau 'mode'. Landau modes are not members of the complete set of true modes and hence are linear combinations of true modes.

A star cluster has many more true modes than the equivalent gas ball because much more information is required to specify a DF than to specify

the density and pressure in a ball of gas. For this reason one hesitates to enumerate a cluster's van Kampen modes, except possibly in the simplified case of vanishing self-gravity. The third contribution to this series (not included in the thesis) shows that we can avoid this enumeration by showing how to decompose any initial state $F(\mathbf{w})$ of the system into a linear combination $F = \int \mathrm{d}\omega\, f(\omega)$ of van Kampen DFs $f(\mathbf{w}, \omega)$. This decomposition automatically identifies the particular mode at frequency ω that is required to synthesise the given DF.

When a model cluster is realised, its van Kampen modes acquire non-zero amplitudes by virtue of Poisson noise. The phases of modes evolve on a dynamical timescale while their amplitudes probably evolve on the two-body timescale. Consequently, there is an early phase of relaxation in the evolution of a simulated cluster in which system-scale distortions emerge. Consideration of the way modes depend of the degree of self-gravity explains why system-scale distortions are less prominent in simulations that are less self-gravitating.

The positivity of the energies of van Kampen modes opens the door to the application of standard statistical physics to stellar systems: while in the long term systems will drift through core collapse and evaporation to states of ever higher entropy, in the medium term disturbed systems may relax to distributions of energy among van Kampen modes that maximise entropy.

2.9 Appendix: Matrix elements of Antonov's operator

Here we compute a general matrix element $\langle \widetilde{f}_- | K | f_- \rangle$ of Antonov's operator, with f and $\widetilde{f}$ any two DFs. Equation (2.23) yields

$$\langle \widetilde{f}_- | K | f_- \rangle = \int \frac{\mathrm{d}^6\mathbf{w}}{f_0'} \widetilde{f}_-^* \left(\mathbf{\Omega} \cdot \frac{\partial}{\partial \boldsymbol{\theta}} \right)^2 f_-$$
$$+ \mathrm{i}\frac{\xi}{\mathcal{E}} \int \mathrm{d}^6\mathbf{w}\, \widetilde{f}_-^* \mathbf{\Omega} \cdot \frac{\partial}{\partial \boldsymbol{\theta}} \sum_\alpha \Phi^{(\alpha)}(\mathbf{x}) j_\alpha[f_1]. \tag{2.55}$$

Since we can write $d^6\mathbf{w} = d^3\mathbf{J}d^3\boldsymbol{\theta}$, we can shift the derivatives wrt $\boldsymbol{\theta}$ around by partial integration and obtain

$$
\begin{aligned}
\langle \widetilde{f}_- | K | f_- \rangle = & -\int \frac{d^6\mathbf{w}}{f_0'} \left(\boldsymbol{\Omega} \cdot \frac{\partial \widetilde{f}_-^*}{\partial \boldsymbol{\theta}} \right) \left(\boldsymbol{\Omega} \cdot \frac{\partial f_-}{\partial \boldsymbol{\theta}} \right) \\
& - i \int d^6\mathbf{w} \sum_\alpha \Phi^{(\alpha)}(\mathbf{x}) \boldsymbol{\Omega} \cdot \frac{\partial \widetilde{f}_-^*}{\partial \boldsymbol{\theta}} j_\alpha[f_1] \\
= & -\int \frac{d^6\mathbf{w}}{f_0'} \left(\boldsymbol{\Omega} \cdot \frac{\partial \widetilde{f}_-^*}{\partial \boldsymbol{\theta}} \right) \left(\boldsymbol{\Omega} \cdot \frac{\partial f_-}{\partial \boldsymbol{\theta}} \right) \\
& - i \frac{\mathcal{E}}{\xi} \sum_\alpha j_\alpha^*[\widetilde{f}_1] j_\alpha[f_1].
\end{aligned}
\tag{2.56}
$$

The symmetry of the rhs wrt $f, \widetilde{f}$ implies that K is Hermitian. Since K is Hermitian, all its eigenvalues ω^2 are real and K's eigenfunctions are either sinusoidal or show pure exponential growth/decay.

When we set $\widetilde{f} = f$, we obtain

$$
\langle f_- | K | f_- \rangle = \int \frac{d^6\mathbf{w}}{|f_0'|} \left| \boldsymbol{\Omega} \cdot \frac{\partial f_-}{\partial \boldsymbol{\theta}} \right|^2 - \frac{\mathcal{E}}{\xi} \sum_\alpha \left| j_\alpha[f_1] \right|^2.
\tag{2.57}
$$

It is interesting to express the rhs of equation (2.57) in terms of f_+ using equations (2.14), (2.19) and (2.25). The result is

$$
\langle f_- | K | f_- \rangle = \int \frac{d^6\mathbf{w}}{|f_0'|} \left| \frac{\partial f_+}{\partial t} \right|^2 - \frac{\mathcal{E}}{\xi} \sum_\alpha \left| \frac{\partial A_\alpha[f_+]}{\partial t} \right|^2.
\tag{2.58}
$$

In the case that f_- is an eigenfunction of K with eigenvalue ω^2, we can replace time derivatives by $-i\omega$ (with ω potentially pure imaginary), and equation (2.58) becomes

$$
\begin{aligned}
\langle f_- | K | f_- \rangle = & \, \omega^2 \langle f_- | f_- \rangle \\
= & \, |\omega^2| \left\{ \int \frac{d^6\mathbf{w}}{|f_0'|} |f_+|^2 - \mathcal{E}_\xi \sum_\alpha \left| A_\alpha[f_+] \right|^2 \right\} \\
= & \, |\omega^2| \left\{ \langle f_+ | f_+ \rangle - \mathcal{E}_\xi \sum_\alpha \left| A_\alpha[f_+] \right|^2 \right\}.
\end{aligned}
\tag{2.59}
$$

2.10 Appendix: Inner product of van Kampen modes

Here we compute the inner product of two van Kampen modes of the same frequency. We have

$$\langle f_- | \widetilde{f}_- \rangle = -\int \frac{\mathrm{d}^6 \mathbf{w}}{f_0'} f_-^* \widetilde{f}_-$$
$$= -(2\pi)^3 \int \frac{\mathrm{d}^3 \mathbf{J}}{f_0'} \sum_{\mathbf{n}} f_{\mathbf{n}-}^* \widetilde{f}_{\mathbf{n}-}. \tag{2.60}$$

Hence $\langle f_- | \widetilde{f}_- \rangle = 0$ unless there is some vector $\mathbf{n}$ at which both $f_{\mathbf{n}-}$ and $\widetilde{f}_{\mathbf{n}-}$ are non-zero. This fact confirms that in the absence of self-gravity, when $f_{\mathbf{n}-}$ vanishes if $g_{\mathbf{n}}$ vanishes, the sought-after basis modes for a given frequency are indexed by $\mathbf{n}$ in the sense that the Fourier expansions of their DFs contain only $\pm\mathbf{n}$. When $\xi > 0$, $f_{\mathbf{n}'-}$ is expected to be non-zero when $g_{\mathbf{n}'} = 0$ providing $g_{\mathbf{n}} \neq 0$ for a vector $\mathbf{n}$ such that $n_3 = n_3'$.

Equation (2.40) gives f_+ for a van Kampen mode, and we have seen that in the case of a mode $|\mathbf{n} \cdot \mathbf{\Omega}| f_{\mathbf{n}-} = \omega f_{\mathbf{n}+}$, so

$$\langle f_- | \widetilde{f}_- \rangle = -(2\pi)^3 \int \frac{\mathrm{d}^3 \mathbf{J}}{f_0'} \sum_{n_1, n_2} \frac{\omega \widetilde{\omega}}{(\mathbf{n} \cdot \mathbf{\Omega})^2} f_{\mathbf{n}+}^* \widetilde{f}_{\mathbf{n}+}$$
$$= -(2\pi)^3 \omega \widetilde{\omega} \, \mathcal{P} \int \frac{\mathrm{d}^3 \mathbf{J}}{f_0'} \sum_{n_1, n_2} \frac{1}{(\mathbf{n} \cdot \mathbf{\Omega})^2}$$
$$\times \left(\frac{(\mathbf{n} \cdot \mathbf{\Omega})^2 f_0'}{\omega^2 - (\mathbf{n} \cdot \mathbf{\Omega})^2} \sum_\alpha A_\alpha^* \Phi_{\mathbf{n}}^{(\alpha)*}(\mathbf{J}) + g_{\mathbf{n}}^*(\mathbf{J}) \delta(\omega^2 - (\mathbf{n} \cdot \mathbf{\Omega})^2) \right)$$
$$\times \left(\frac{(\mathbf{n} \cdot \mathbf{\Omega})^2 f_0'}{\widetilde{\omega}^2 - (\mathbf{n} \cdot \mathbf{\Omega})^2} \sum_\alpha \widetilde{A}_\alpha \Phi_{\mathbf{n}}^{(\alpha)}(\mathbf{J}) + \widetilde{g}_{\mathbf{n}}(\mathbf{J}) \delta(\widetilde{\omega}^2 - (\mathbf{n} \cdot \mathbf{\Omega})^2) \right),$$

$$\tag{2.61}$$

where quantities associated with the mode $\widetilde{f}$ are marked by tildes. When we multiply out the big brackets, we get a term with two denominators of the form $(\mathbf{n} \cdot \mathbf{\Omega})^2 - \omega^2$. The integral over these is to be interpreted as a principal value, that is by excluding points at which the denominator vanishes. We use

the identity (e.g. Ramos and White 2018)

$$\mathcal{P}\frac{1}{x-x_1}\mathcal{P}\frac{1}{x-x_2} = \mathcal{P}\frac{1}{x_1-x_2}\left(\mathcal{P}\frac{1}{x-x_1} - \mathcal{P}\frac{1}{x-x_2}\right)$$
$$+ \pi^2\delta(x-x_1)\delta(x_1-x_2) \tag{2.62}$$

to rewrite this term as

$$\int \mathrm{d}^3\mathbf{J}\cdots =\mathcal{P}\int \frac{\mathrm{d}^3\mathbf{J}}{f_0'} \sum_{n_1,n_2}(\mathbf{n}\cdot\mathbf{\Omega}f_0')^2 \sum_{\alpha\beta}A_\alpha^*\widetilde{A}_\beta\Phi_\mathbf{n}^{(\alpha)*}\Phi_\mathbf{n}^{(\beta)}$$
$$\times \left\{\frac{1}{\omega^2-\widetilde{\omega}^2}\left(\frac{1}{(\mathbf{n}\cdot\mathbf{\Omega})^2-\omega^2} - \frac{1}{(\mathbf{n}\cdot\mathbf{\Omega})^2-\widetilde{\omega}^2}\right)\right.$$
$$\left.+ \pi^2\delta((\mathbf{n}\cdot\mathbf{\Omega})^2-\omega^2)\delta(\omega^2-\widetilde{\omega}^2)\right\}. \tag{2.63}$$

The cross terms in the product of equation (2.61) can be written

$$\int \mathrm{d}^3\mathbf{J}\cdots = \int \frac{\mathrm{d}^3\mathbf{J}}{f_0'}\frac{f_0'}{\omega^2-\widetilde{\omega}^2}\sum_{n_1,n_2}$$
$$\times \left(\sum_\alpha A_\alpha^*\Phi_\mathbf{n}^{(\alpha)*}\widetilde{g}_\mathbf{n}\delta((\mathbf{n}\cdot\mathbf{\Omega})^2-\widetilde{\omega}^2)\right.$$
$$\left.-\sum_\beta \widetilde{A}_\beta\Phi_\mathbf{n}^{(\beta)}g_\mathbf{n}^*\delta((\mathbf{n}\cdot\mathbf{\Omega})^2-\omega^2)\right), \tag{2.64}$$

while the term involving $g_\mathbf{n}^*\widetilde{g}_\mathbf{n}$ can be written

$$\int \mathrm{d}^3\mathbf{J}\cdots = \int \frac{\mathrm{d}^3\mathbf{J}}{f_0'}\sum_{n_1,n_2}\frac{g_\mathbf{n}^*\widetilde{g}_\mathbf{n}}{(\mathbf{n}\cdot\mathbf{\Omega})^2}\,\delta((\mathbf{n}\cdot\mathbf{\Omega})^2-\omega^2)\,\delta(\omega^2-\widetilde{\omega}^2). \tag{2.65}$$

Adding these fragments together and reinstating the prefactor in equation

(2.61) we have

$$
\begin{aligned}
\langle f_- | \widetilde{f}_- \rangle = {}&-(2\pi^3)\omega\widetilde{\omega}\,\mathcal{P}\!\int \frac{\mathrm{d}^3\mathbf{J}}{f_0'} \sum_{n_1,n_2} \left\{ \frac{f_0'}{\omega^2 - \widetilde{\omega}^2} \right[\\
&\times \left(\frac{(\mathbf{n}\cdot\boldsymbol{\Omega})^2 f_0'}{(\mathbf{n}\cdot\boldsymbol{\Omega})^2 - \omega^2} \sum_\alpha A_\alpha^* \Phi_\mathbf{n}^{(\alpha)*} + g_\mathbf{n}^* \delta((\mathbf{n}\cdot\boldsymbol{\Omega})^2 - \omega^2) \right) \sum_\beta \widetilde{A}_\beta \Phi_\mathbf{n}^{(\beta)} \\
&- \left(\frac{(\mathbf{n}\cdot\boldsymbol{\Omega})^2 f_0'}{(\mathbf{n}\cdot\boldsymbol{\Omega})^2 - \widetilde{\omega}^2} \sum_\beta \widetilde{A}_\beta \Phi_\mathbf{n}^{(\beta)} + \widetilde{g}\delta((\mathbf{n}\cdot\boldsymbol{\Omega})^2 - \widetilde{\omega}^2) \right) \sum_\alpha A_\alpha^* \Phi_\mathbf{n}^{(\alpha)*} \right] \\
&+ \left(\pi^2(\mathbf{n}\cdot\boldsymbol{\Omega}f_0')^2 \sum_{\alpha\beta} A_\alpha^* \widetilde{A}_\beta \Phi_\mathbf{n}^{(\alpha)*}\Phi_\mathbf{n}^{(\beta)} + \frac{g_\mathbf{n}^*\widetilde{g}_\mathbf{n}}{(\mathbf{n}\cdot\boldsymbol{\Omega})^2} \right) \\
&\times \delta((\mathbf{n}\cdot\boldsymbol{\Omega})^2 - \omega^2)\,\delta(\omega^2 - \widetilde{\omega}^2) \bigg\} \\
= {}&-(2\pi)^3 \frac{\omega\widetilde{\omega}}{\omega^2 - \widetilde{\omega}^2} \int \mathrm{d}^3\mathbf{J} \sum_{n_1,n_2} \left(f_{\mathbf{n}+}^* \sum_\beta \widetilde{A}_\beta \Phi_\mathbf{n}^{(\beta)} \right. \\
&\left. - \widetilde{f}_{\mathbf{n}+} \sum_\alpha A_\alpha^* \Phi_\mathbf{n}^{(\alpha)*} \right) \\
&-(2\pi)^3 \omega\widetilde{\omega} \int \mathrm{d}^3\mathbf{J} \sum_{n_1,n_2} \left(\pi^2(\mathbf{n}\cdot\boldsymbol{\Omega})^2 f_0' \sum_{\alpha\beta} A_\alpha^* \widetilde{A}_\beta \Phi_\mathbf{n}^{(\alpha)*}\Phi_\mathbf{n}^{(\beta)} \right. \\
&\left. + \frac{g_\mathbf{n}^*\widetilde{g}_\mathbf{n}}{(\mathbf{n}\cdot\boldsymbol{\Omega})^2 f_0'} \right) \delta((\mathbf{n}\cdot\boldsymbol{\Omega})^2 - \omega^2)\,\delta(\omega^2 - \widetilde{\omega}^2). \qquad (2.66)
\end{aligned}
$$

Now

$$
\begin{aligned}
A_\alpha &= \int \mathrm{d}^6\mathbf{w}\, \Phi^{(\alpha)*} f_+ \\
&= \int \mathrm{d}^3\mathbf{J}\,\mathrm{d}^3\boldsymbol{\theta} \sum_\mathbf{n} \Phi_\mathbf{n}^{(\alpha)*} \mathrm{e}^{-\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}} f_+ = \sum_\mathbf{n} \Phi_\mathbf{n}^{(\alpha)*} f_{\mathbf{n}+}, \qquad (2.67)
\end{aligned}
$$

so the first integral in our final expression for $\langle f_- | f_- \rangle$ vanishes because its integrand is $\sum_\alpha (\widetilde{A}_\alpha A_\alpha^* - A_\alpha^* \widetilde{A}_\alpha)$. With some further simplifications we can write the inner product in the form given by equation (2.52).

Chapter 3

Modes of a Stellar System II: non-ergodic systems[1]

Abstract

An equation is derived for the energy of a small disturbance in a system that is generated by a distribution function (DF) of the form $f(\mathbf{J})$ – most galaxies and star clusters can be closely approximated by such a DF. The theory of van Kampen modes is extended to such general systems. A bilinear form on the space of DFs is defined such that the energy of a disturbance is its norm under this form. It is shown that van Kampen modes that differ in frequency are then orthogonal, with the consequence that the energies of van Kampen modes are additive. Consequently, most of the insight into the dynamics of ergodic systems that was gained in the previous chapter on the van Kampen modes of ergodic systems applies to real clusters and galaxies.

3.1   Introduction

Stellar systems have hitherto been modelled in mean-field limit of an infinite number of constituent particles, when fluctuations vanish. That fluctuations play an essential role in the evolution of clusters was recognised over half a century ago (Henon, Spitzer, Chandra), but even now observations are fitted to mean-field models such as Michie-King models (Michie 1963, King 1966, McLaughlin et al. 2006, Piotto et al. 2015, Claydon et al. 2019).

Some fluctuations are internally generated by the shot noise inherent in a system with a finite number of particles, while other fluctuations are externally stimulated by the gravitational fields of neighbouring systems. As observational data become more precise, there must come a point at which fluctuations of either type can be detected. Detection of fluctuations would

open the possibility of using galaxies and star clusters to detect the passage of dark-matter haloes because the tidal fields of such haloes will excite the large-scale modes of globular clusters and dwarf spheroidal galaxies.

Modelling the small-scale effects of shot noise is rather straightforward because the system's self-gravity is only important on the largest scales (e.g Fouvry et al. 2021). But following the derivation of the Balescu-Lenard equation by Heyvaerts (2010) and Chavanis (2012), it has become evident that self-gravity has a big impact on system-scale fluctuations even when the latter are stimulated by Poisson noise (Fouvry et al. 2015, Hamilton et al. 2018, Lau and Binney 2019; 2021c, Heggie et al. 2020). One expects externally generated fluctuations to be predominantly large-scale, so a viable theory of fluctuations must encompass self-gravity.

The standard approach to a theory of fluctuations is via normal modes: By linearising the equations of motion one derives a set of harmonically oscillating disturbances that is complete in the sense that any initial condition can be expressed as a linear combination of modes. This decomposition simultaneously characterises the initial condition in a physically significant way, and provides a convenient way to compute the system's evolution by taking advantage of the simple rule for evolving a normal mode: that of a sinusoidal time-dependence.

The stability of stellar systems has traditionally been investigated by determining the frequencies of 'Landau modes' because the system is unstable if any of these frequencies has a positive imaginary part. Landau modes, however, lack the essential completeness property of normal modes. In the case of a homogeneous electrostatic plasma, van Kampen (1955) presented the true normal modes, now known as van Kampen modes in his honour, and Case (1959) proved that van Kampen's modes are complete. Specification of the distribution function (DF) $f(x, v)$ of a collisionless system requires much more information than is required to specify the state of a fluid, and this fact is reflected in key differences between van Kampen modes and the modes of a

fluid system. Crucially, the spectrum of van Kampen modes contains a continuum of real frequencies, and the DFs of individual modes contains a Dirac δ-function associated with its frequency. Expression of an initial condition as a sum of normal modes involves an integral over frequency that eliminates the δ-function.

In Chapter 2 of this series (Lau and Binney 2021a) we derived the van Kampen modes of ergodic stellar systems, that is ones with DFs $f(H)$, where $H = \frac{1}{2}v^2 + \Phi(\mathbf{x})$ is the Hamiltonian of an individual star. They showed that the energy of a disturbance is the sum of the energies of its constituent van Kampen modes, and that when the system is stable the energies of all modes are inherently positive. In this chapter we generalise these results to non-ergodic stellar systems, that is, systems with DF $f(\mathbf{J})$, where $\mathbf{J}$ is the vector exhibit velocity anisotropy and/or systematic rotation. Some our results will be restricted to the subclass of systems that are unchanged by reversing all velocities, which in practice excludes systems with systematic rotation.

In Section 3.2 we introduce the notation and some results that will be required in later sections. In Section 3.3 we extend to systems with DF $f(\mathbf{J})$ the formula for the energy of a fluctuation that Nelson and Tremaine (1999) derived for ergodic systems. In Section 3.4 we obtain results for the van Kampen modes of systems with a DF $f(\mathbf{J})$ that are analogues of results derived for ergodic systems in Chapter 2. The route taken to these results does, however, differ significantly to that used in Chapter 2, being simpler but less powerful – the fourth contribution (not included in this thesis) in this series will present the rather intricate generalisation to non-ergodic systems of the approach used in Chapter 2, which exploits the Hermitian operator introduced by Antonov (1961). In Section 3.5 we discuss the differences between ergodic and non-ergodic systems and the relation between true modes and Landau modes.

3.2 Mathematical background

Here we introduce essential mathematical tools and establish our notation. We focus on systems with integrable mean-field potentials, so their DFs can be written in the form $f(\mathbf{J})$.

3.2.1 Angle-action variables

The role that Cartesian variables play for homogeneous systems is played for spheroidal systems by angle-action variables $(\boldsymbol{\theta}, \mathbf{J})$. The actions J_i are constants of motion while their conjugate variables, the angles θ_i, increase linearly in time, so $\boldsymbol{\theta}(t) = \boldsymbol{\theta}(0) + \boldsymbol{\Omega}t$. The particles' Hamiltonian $H(\mathbf{x}, \mathbf{v})$ is a function $H(\mathbf{J})$ of the actions only and the frequencies Ω_i that control the rates of increase of the angles are given by $\boldsymbol{\Omega} = \partial H/\partial \mathbf{J}$. Angle-action variables are canonical, so the volume element of phase space $\mathrm{d}^6\mathbf{w} = \mathrm{d}^3\mathbf{x}\mathrm{d}^3\mathbf{v} = \mathrm{d}^3\boldsymbol{\theta}\mathrm{d}^3\mathbf{J}$ and Poisson brackets can be computed as

$$[f, g] = \sum_i \left(\frac{\partial f}{\partial \theta_i} \frac{\partial g}{\partial J_i} - \frac{\partial f}{\partial J_i} \frac{\partial g}{\partial \theta_i} \right). \tag{3.1}$$

Functions on phase space can be expressed as Fourier series:

$$h(\mathbf{w}) = \sum_{\mathbf{n}} h_{\mathbf{n}}(\mathbf{J}) \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}} \; ; \; h_{\mathbf{n}}(\mathbf{J}) = \int \frac{\mathrm{d}^3\boldsymbol{\theta}}{(2\pi)^3} \mathrm{e}^{-\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}} h(\mathbf{w}). \tag{3.2}$$

Note that for real h, $h_{-\mathbf{n}} = h_{\mathbf{n}}^*$.

3.2.2 Potential-density pairs

Following Kalnajs (1976) we solve Poisson's equation by introducing a basis of biorthogonal potential-density pairs. That is, a set of pairs $(\rho^{(\alpha)}, \Phi^{(\alpha)})$ such that

$$4\pi G\rho^{(\alpha)} = \nabla^2\Phi^{(\alpha)} \quad \text{and} \quad \int \mathrm{d}^3\mathbf{x}\, \Phi^{(\alpha)*}\rho^{(\alpha')} = -\mathcal{E}\delta_{\alpha\alpha'}, \tag{3.3}$$

where $\mathcal{E}$ is an arbitrary constant with the dimensions of energy. Given a density distribution $\rho(\mathbf{x})$, we expand it in the basis

$$\rho(\mathbf{x}) = \sum_\alpha A_\alpha \rho^{(\alpha)}(\mathbf{x}) \quad \Leftrightarrow \quad \Phi(\mathbf{x}) = \sum_\alpha A_\alpha \Phi^{(\alpha)}(\mathbf{x}), \tag{3.4}$$

where

$$A_\alpha = -\frac{1}{\mathcal{E}} \int d^6\mathbf{w}\, \Phi^{(\alpha)*}(\mathbf{x}) f(\mathbf{w}). \tag{3.5}$$

If ρ and Φ are time-dependent, the A_α become time-dependent. From equations (3.3) and (3.4) one can obtain an expression for Φ in terms of ρ. Comparison of this relation with Poisson's integral, yields

$$\frac{G}{|\mathbf{x}' - \mathbf{x}|} = \frac{1}{\mathcal{E}} \sum_\alpha \Phi^{(\alpha)}(\mathbf{x}) \Phi^{(\alpha)*}(\mathbf{x}'). \tag{3.6}$$

The system's potential energy is

$$P = \tfrac{1}{2} \int d^3\mathbf{x}\, \Phi\rho = \tfrac{1}{2} \int d^3\mathbf{x} \sum_\alpha A_\alpha \Phi^{(\alpha)} \rho$$
$$= -\frac{\mathcal{E}}{2} \sum_\alpha |A_\alpha|^2. \tag{3.7}$$

A related calculation is the potential energy of one system when placed in the potential of another. This is

$$P' = -G \int d^3\mathbf{x}\, \frac{\rho(\mathbf{x})\rho'(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}. \tag{3.8}$$

The symmetry of this expression establishes that the energy of system a in the potential of system b is the same as that of system b placed in the potential of system a. Adaptation of the derivation of equation (3.7) shows that

$$P' = -\mathcal{E} \sum_\alpha A_\alpha^* A_\alpha'. \tag{3.9}$$

Now

$$P' = \int d^6\mathbf{w}\, f\Phi' = \int d^3\mathbf{J}\, d^3\boldsymbol{\theta} \sum_{\mathbf{mn}} f_\mathbf{m} \Phi'_\mathbf{n} e^{i(\mathbf{m}+\mathbf{n})\cdot\boldsymbol{\theta}}$$
$$= (2\pi)^3 \sum_\mathbf{n} \int d^3\mathbf{J}\, f_\mathbf{n}^* \Phi'_\mathbf{n}. \tag{3.10}$$

3.2.3 Linearised CBE

On dynamical timescales the DF of a stellar system satisfies the collisionless Boltzmann equation

$$\frac{\partial f}{\partial t} + [f, H] = 0. \tag{3.11}$$

When we split $f(\boldsymbol{\theta}, \mathbf{J}, t) = f_0(\mathbf{J}) + f_1(\boldsymbol{\theta}, \mathbf{J}, t)$ into its mean-field and fluctuating components and neglect terms quadratic and higher in the fluctuations, the CBE can be written

$$\frac{\partial f_1}{\partial t} + [f_1, H_0] + [f_0, \Phi_1] = 0. \tag{3.12}$$

When we use angle-action coordinates to evaluate the Poisson brackets and write f_1 and Φ_1 as Fourier series in angles, equation (3.12) yields

$$\frac{\partial f_\mathbf{n}}{\partial t} + \mathrm{i}\mathbf{n} \cdot \boldsymbol{\Omega} f_\mathbf{n} - \mathrm{i}\mathbf{n} \cdot \frac{\partial f_0}{\partial \mathbf{J}} \Phi_{1\mathbf{n}} = 0, \tag{3.13}$$

where we have dropped the subscript 1 from $f_{1\mathbf{n}}$ for brevity but retained it on $\Phi_{1\mathbf{n}}$ for reasons that will soon become apparent.

3.3 Energy of a perturbation

Following Nelson and Tremaine (1999) we imagine using an externally applied gravitational field to impose a real perturbation f_1 on a mean-field model $f_0(\mathbf{J})$. The perturbing gravitational potential Φ_1 now has two components, the potential Φ_e of the externally applied field and the potential Φ generated via Poisson's equation by the perturbed density distribution

$$\rho_1 = \int \mathrm{d}^3\mathbf{v} \, f_1. \tag{3.14}$$

The rate at which work is done by the external field is

$$\begin{aligned} \frac{\mathrm{d}E}{\mathrm{d}t} &= -\int \mathrm{d}^3\mathbf{x} \, \mathbf{v} \cdot \frac{\partial \Phi_\mathrm{e}}{\partial \mathbf{x}} \rho_1 \\ &= -\int \mathrm{d}^6\mathbf{w} \, f_1(\mathbf{w})\mathbf{v} \cdot \frac{\partial \Phi_\mathrm{e}}{\partial \mathbf{x}}. \end{aligned} \tag{3.15}$$

Now

$$\mathbf{v} \cdot \frac{\partial \Phi_\mathrm{e}}{\partial \mathbf{x}} = -[H_0, \Phi_\mathrm{e}] = \boldsymbol{\Omega} \cdot \frac{\partial \Phi_\mathrm{e}}{\partial \boldsymbol{\theta}}, \tag{3.16}$$

so equation (3.15) can be written

$$\begin{aligned} \frac{\mathrm{d}E}{\mathrm{d}t} &= -\int \mathrm{d}^6\mathbf{w} \, f_1(\mathbf{w})\boldsymbol{\Omega} \cdot \frac{\partial \Phi_\mathrm{e}}{\partial \boldsymbol{\theta}} \\ &= -(2\pi)^3 \int \mathrm{d}^3\mathbf{J} \sum_\mathbf{n} f_\mathbf{n}^*(\mathbf{w})\mathrm{i}\boldsymbol{\Omega} \cdot \mathbf{n} \, \Phi_{\mathrm{e}\mathbf{n}}. \end{aligned} \tag{3.17}$$

With Φ_1 decomposed into its two components, the linearised CBE (3.13) can be written

$$\frac{\partial f_{\mathbf{n}}}{\partial t} + i\mathbf{n} \cdot \mathbf{\Omega} f_{\mathbf{n}} - i\mathbf{n} \cdot \frac{\partial f_0}{\partial \mathbf{J}} \Phi_{s\mathbf{n}} = i\mathbf{n} \cdot \frac{\partial f_0}{\partial \mathbf{J}} \Phi_{e\mathbf{n}}. \tag{3.18}$$

Eliminating $\Phi_{e\mathbf{n}}$ between the last two equations

$$\frac{\mathrm{d}E}{\mathrm{d}t} \equiv -(2\pi)^3 \int \mathrm{d}^3\mathbf{J} \sum_{\mathbf{n}} f_{\mathbf{n}}^* \frac{\mathbf{n} \cdot \mathbf{\Omega}}{\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0}$$
$$\times \left(\frac{\partial f_{\mathbf{n}}}{\partial t} + i\mathbf{n} \cdot \mathbf{\Omega} f_{\mathbf{n}} - i\mathbf{n} \cdot \frac{\partial f_0}{\partial \mathbf{J}} \Phi_{\mathbf{n}} \right) \tag{3.19}$$

The integrand contains three terms. The middle term is proportional to $\mathbf{n} \cdot \mathbf{\Omega}|f_{\mathbf{n}}|^2$ and vanishes when summed over $\mathbf{n}$ because the sum includes both $\mathbf{n}$ and $-\mathbf{n}$. The integral over the first term yields

$$\frac{\mathrm{d}K}{\mathrm{d}t} = -\frac{(2\pi)^3}{2} \int \mathrm{d}^3\mathbf{J} \sum_{\mathbf{n}} \frac{\mathbf{n} \cdot \mathbf{\Omega}}{\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0} \frac{\partial |f_{\mathbf{n}}|^2}{\partial t}. \tag{3.20}$$

In preparation for handling the third term, we note that

$$\int \mathrm{d}^3\boldsymbol{\theta} \, f_1[\Phi, H_0] = \int \mathrm{d}^3\boldsymbol{\theta} \, f_1 \frac{\partial \Phi_s}{\partial \boldsymbol{\theta}} \cdot \mathbf{\Omega}$$
$$= (2\pi)^3 \sum_{\mathbf{n}} f_{\mathbf{n}}^* i\mathbf{n} \cdot \mathbf{\Omega} \Phi_{\mathbf{n}}. \tag{3.21}$$

Hence the third term in equation (3.19) is

$$\frac{\mathrm{d}P}{\mathrm{d}t} \equiv (2\pi)^3 \int \mathrm{d}^3\mathbf{J} \sum_{\mathbf{n}} i\mathbf{n} \cdot \mathbf{\Omega} f_{\mathbf{n}}^* \Phi_{\mathbf{n}}$$
$$= \int \mathrm{d}^6\mathbf{w} \, f_1[\Phi, H_0] = \int \mathrm{d}^6\mathbf{w} \, f_1 \mathbf{v} \cdot \partial \Phi_{\overline{\frac{\partial}{\partial \mathbf{x}}}}$$
$$= \int \mathrm{d}^3\mathbf{x} \, \Phi_{\mathbf{s}} \frac{\partial \rho_1}{\partial t} = -\frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{2} G \int \mathrm{d}^3\mathbf{x} \, \mathrm{d}^3\mathbf{x}' \frac{\rho_1(\mathbf{x})\rho_1(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \tag{3.22}$$

where the penultimate equality uses integration by parts and the continuity equation $\partial \rho / \partial t = -\nabla_{\mathbf{x}} \cdot (\rho \mathbf{v})$.

Now that the two surviving contributions to the integral in equation (3.19) have proved to be total time derivatives, we can immediately integrate from

$E = K = P = 0$ at $t = 0$ to obtain an equation for the energy of an arbitrary fluctuation

$$E = -\frac{(2\pi)^3}{2} \int \mathrm{d}^3\mathbf{J} \sum_{\mathbf{n}} \frac{\mathbf{n}\cdot\boldsymbol{\Omega}}{\mathbf{n}\cdot\nabla_{\mathbf{J}}f_0}|f_{\mathbf{n}}|^2 - \tfrac{1}{2} \int \mathrm{d}^3\mathbf{x}\,\mathrm{d}^3\mathbf{x}' \frac{G\rho_1^2}{|\mathbf{x}-\mathbf{x}'|}. \qquad (3.23)$$

This equation generalises to three dimensions the quantity Kalnajs (1971) shows to be constant in an isolated razor-thin disc (Kalnajs' eqn. 48), and which he says he will 'identify as energy'. Our derivation shows that E is the work that must be done to establish a perturbation rather than just showing that E is constant when Φ_{e} vanishes. Equation (3.23) differs from equation (5.130) in Binney and Tremaine (2008) for the energy of a disturbed ergodic system only by the replacement of $\mathrm{d}f_0/\mathrm{d}H$ by $\mathbf{n}\cdot\nabla_{\mathbf{J}}f_0/\mathbf{n}\cdot\boldsymbol{\Omega}$. When equation (3.6) is used to eliminate $|\mathbf{x}-\mathbf{x}'|$ from equation (3.23) we obtain

$$E = -\frac{(2\pi)^3}{2} \int \mathrm{d}^3\mathbf{J} \sum_{\mathbf{n}} \frac{\mathbf{n}\cdot\boldsymbol{\Omega}}{\mathbf{n}\cdot\nabla_{\mathbf{J}}f_0}|f_{\mathbf{n}}|^2 - \tfrac{1}{2}\mathcal{E} \sum_{\alpha} |A_{\alpha}|^2. \qquad (3.24)$$

3.3.1 Restriction of the perturbed DF

For certain combinations of $\mathbf{n}$ and $\mathbf{J}$, $\mathbf{n}\cdot\nabla_{\mathbf{J}}f_0$ will vanish. On first sight, it seems as if such combinations will make the integral in equation (3.23) for E ill defined. However, from equation (3.18) it follows that for these combinations $f_{\mathbf{n}}(\mathbf{J})$ remains zero as the disturbance is excited, so the integrand in equation (3.23) vanishes at the apparently problematic points. Thus the disturbances f that can be induced in a system with equilibrium DF $f_0(\mathbf{J})$ by an external potential are restricted in form.

Antonov (1961) broadened this result by showing (e.g. Binney and Tremaine 2008, p. 429) that the the perturbation f_1 that is generated by applying *any* disturbing Hamiltonian H_1 to f_0 will be of the form

$$f_1 = [h, f_0] = \frac{\partial h}{\partial \boldsymbol{\theta}} \cdot \frac{\partial f_0}{\partial \mathbf{J}}, \qquad (3.25)$$

where $h(\mathbf{w})$ is a function that depends on H_1. On expanding f and h in Fourier series we obtain

$$f_{\mathbf{n}} = \mathrm{i}\mathbf{n} \cdot \frac{\partial f_0}{\partial \mathbf{J}} h_{\mathbf{n}} \qquad (3.26)$$

so when $\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0 = 0$, $f_{\mathbf{n}}$ vanishes no matter how the system is disturbed.

3.3.2 A bilinear form for DFs

Additivity of energies is a fundamental property of normal modes. It emerges if we can express the energy of a disturbance as the norm of the disturbance defined by a bilinear form under which normal modes are mutually orthogonal. Equation (3.24) suggests that the required form is

$$\langle f | \widetilde{f} \rangle = -\frac{(2\pi)^3}{2} \int \mathrm{d}^3 \mathbf{J} \sum_{\mathbf{n}} \frac{\mathbf{n} \cdot \mathbf{\Omega}}{\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0} f_{\mathbf{n}}^* \widetilde{f}_{\mathbf{n}} - \tfrac{1}{2} \mathcal{E} \sum_{\alpha} A_{\alpha}^* A_{\alpha}', \tag{3.27}$$

where $A_{\alpha}[f]$ is a functional of f while A_{α}' denotes the corresponding functional of $\widetilde{f}$.

3.4 Normal modes of a non-ergodic system

We now look for disturbances with exponential time dependence, so

$$f_{\mathbf{n}}(\mathbf{J}, t) = f_{\mathbf{n}}(\mathbf{J}, \omega) \mathrm{e}^{-\mathrm{i}\omega t}, \quad \Phi_{\mathbf{n}}(\mathbf{J}, t) = \Phi_{\mathbf{n}}(\mathbf{J}, \omega) \mathrm{e}^{-\mathrm{i}\omega t} \tag{3.28}$$

With this ansatz, the linearised CBE (3.13) becomes

$$(\mathbf{n} \cdot \mathbf{\Omega} - \omega) f_{\mathbf{n}} = (\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0) \Phi_{\mathbf{n}}(\mathbf{J}, \omega). \tag{3.29}$$

This equation yields a well defined value for $f_{\mathbf{n}}$ in terms of $\Phi_{\mathbf{n}}$ when $\mathbf{n} \cdot \mathbf{\Omega} \neq \omega$, but on resonant tori (tori on which $\mathbf{n} \cdot \mathbf{\Omega} = 0$) it does not constrain $f_{\mathbf{n}}$. Therefore, as van Kampen (1955) pointed out, the solution to this equation must be written

$$f_{\mathbf{n}}(\mathbf{J}, \omega) = \mathcal{P} \frac{\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0}{\mathbf{n} \cdot \mathbf{\Omega} - \omega} \Phi_{\mathbf{n}}(\mathbf{J}, \omega) + g_{\mathbf{n}}(\mathbf{J}) \delta(\mathbf{n} \cdot \mathbf{\Omega} - \omega), \tag{3.30}$$

where $\mathcal{P}$ indicates that when integrated wrt $\mathbf{n} \cdot \mathbf{\Omega}$ or ω, the Cauchy principal value around the singularity at $\omega = \mathbf{n} \cdot \mathbf{\Omega}$ should be taken, and $g_{\mathbf{n}}(\mathbf{J})$ is an arbitrary function that specifies non-trivial distributions of stars on the resonant tori. In Section 3.3.1 we saw that the DFs of disturbances are restricted such than $f_{\mathbf{n}} = 0$ when $\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0 = 0$. Since a system's normal modes comprise possible disturbances, $g_{\mathbf{n}}$ must be restricted in the same way:

$$\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0 = 0 \quad \Rightarrow \quad g_{\mathbf{n}} = 0. \tag{3.31}$$

In Chapter 2 $\Phi_{\mathbf{n}}(\mathbf{J}, \omega)$ in equation (3.30) is interpreted as the potential generated by the the resonant stars and by the non-resonant stars that they excite. It is the *dressed* potential of the resonant stars. Consequently, Poisson's equation imposes a relation between $g_{\mathbf{n}}$ and $\Phi_{\mathbf{n}}$. We obtain this relation by multiplying equation (3.30) by $\mathrm{d}^6\mathbf{w}\, \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}}\Phi^{(\alpha')*}$ and integrating through phase space to obtain the self-consistency condition (cf equation (2.42))

$$\sum_{\alpha} M_{\alpha'\alpha} A_{\alpha} = -B_{\alpha'} \tag{3.32}$$

where

$$M_{\alpha'\alpha}(\omega) = \delta_{\alpha'\alpha} - \frac{(2\pi)^3}{\mathcal{E}}\mathcal{P}\int \mathrm{d}^3\mathbf{J} \sum_{\mathbf{n}} \frac{\mathbf{n}\cdot\nabla f_0}{\mathbf{n}\cdot\boldsymbol{\Omega} - \omega}\Phi_{\mathbf{n}}^{(\alpha')*}\Phi_{\mathbf{n}}^{(\alpha)}$$

$$B_{\alpha'} = -\frac{(2\pi)^3}{\mathcal{E}}\int \mathrm{d}^3\mathbf{J} \sum_{\mathbf{n}} g_{\mathbf{n}}(\mathbf{J})\Phi_{\mathbf{n}}^{(\alpha')*}(\mathbf{J})\delta(\mathbf{n}\cdot\boldsymbol{\Omega} - \omega). \tag{3.33}$$

In any truly stable system (so excluding marginally stable systems such as those discussed by Mathur 1990), $M(\omega)$ has an inverse for all real ω, so given $g(\mathbf{w})$ we can solve for the amplitudes A_{α} of the corresponding density and potential. The coefficients $B_{\alpha'}$ give the density and potential generated by the resonant stars alone, that is, after removing the contributions of stars driven by the gravitational field of the resonant stars (Chapter 2, equation (2.46)).

In principle for given functions $g_{\mathbf{n}}$ equation (3.30) yields a mode for every frequency of the form $\mathbf{n}\cdot\boldsymbol{\Omega}$, but since vectors $\mathbf{n}$ with any large component generate very small densities in real space, the important modes are confined to frequencies that range from near zero ($|\boldsymbol{\Omega}| \approx 0$ at large $|\mathbf{J}|$) to about twice the system's maximum circular frequency.

B necessarily vanishes for complex ω because $\mathbf{n}\cdot\boldsymbol{\Omega}$ is inherently real. Thus all van Kampen modes have real frequencies.

Any additional modes must arise when $|\mathbf{M}(\omega)| = 0$ at some possibly complex frequency ω. If the underlying equilibrium f_0 is unchanged by reversing all velocities (in practice meaning that it has no net rotation), then if $|\mathbf{M}(\omega)| = 0$ then $|\mathbf{M}(-\omega)| = 0$ also, and at one of these two frequencies the disturbance

will grow exponentially and the system is unstable. Hence every mode of a truly stable, time-reversible system is a van Kampen mode. This result extends to flattened systems and systems with velocity anisotropy a key result of Chapter 2. We defer discussion of marginally stable systems to Section 3.5.

The $\mathcal{P}$ and δ symbols in equation (3.30) signal that van Kampen modes make sense only within an integral wrt ω. Hence a physical disturbance will always be of the form

$$F(\boldsymbol{\theta}, \mathbf{J}, t) = \sum_{\mathbf{n}} \int d\omega \, f_{\mathbf{n}}(\mathbf{J}, \omega) e^{i(\mathbf{n}\cdot\boldsymbol{\theta} - \omega t)}. \tag{3.34}$$

This disturbance is specified by one's choice of the functions $g_{\mathbf{n}}(\mathbf{J})$.

3.4.1 Energies of normal modes

In Appendix A we show that in the case of two normal modes $f(\omega)$ and $\widetilde{f}(\widetilde{\omega})$ the bilinear form (3.27) can be brought to the form

$$\langle f|\widetilde{f}\rangle = -\frac{(2\pi)^3}{2} \sum_{\mathbf{n}} \int d^3\mathbf{J} \, \omega \left\{ \pi^2 (\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0) \Phi_{\mathbf{n}}^* \widetilde{\Phi}_{\mathbf{n}} \right.$$
$$\left. + \frac{g_{\mathbf{n}}^* \widetilde{g}_{\mathbf{n}}}{\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0} \right\} \delta(\mathbf{n} \cdot \boldsymbol{\Omega} - \omega) \delta(\omega - \widetilde{\omega}). \tag{3.35}$$

Remarkably, on account of the factor $\delta(\mathbf{n}\cdot\boldsymbol{\Omega} - \omega)$ the bilinear form is computed by integrating only over resonant stars, despite much of the energy lying with driven, non-resonant stars. Combining this equation for the bilinear form evaluated on two van Kampen modes with the equation (3.34) expressing a general disturbance as a sum of van Kampen modes, we obtain an alternative expression for the value of the form on any two disturbances F and $\widetilde{F}$:

$$\langle F|\widetilde{F}\rangle = \left\langle \int d\omega \, f(\omega) \middle| \int d\widetilde{\omega} \, \widetilde{f}(\widetilde{\omega}) \right\rangle = \int d\omega \, d\widetilde{\omega} \, \langle f(\omega)|\widetilde{f}(\widetilde{\omega})\rangle$$
$$= -\frac{(2\pi)^3}{2} \sum_{\mathbf{n}} \int d^3\mathbf{J} \, \omega_{\mathbf{n}}(\mathbf{J})$$
$$\times \left\{ \pi^2 (\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0) \Phi_{\mathbf{n}}^*(\omega_{\mathbf{n}}) \widetilde{\Phi}_{\mathbf{n}}(\omega_{\mathbf{n}}) + \frac{g_{\mathbf{n}}^*(\omega_{\mathbf{n}}) \widetilde{g}_{\mathbf{n}}(\omega_{\mathbf{n}})}{\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0} \right\}, \tag{3.36}$$

where $\omega_{\mathbf{n}}(\mathbf{J}) \equiv \mathbf{n} \cdot \boldsymbol{\Omega}$ and the potentials $\Phi_{\mathbf{n}}$ and driving terms $g_{\mathbf{n}}$ have acquired $\omega_{\mathbf{n}}$ as an argument to indicate that they are the potentials and driving terms of the van Kampen modes with that frequency in the decomposition of F into modes.

When we set $\widetilde{F} = F$ in equation (3.36) we obtain the energy of a general disturbance as the sum of the energies of its component van Kampen modes:

$$E[F] = \langle F|F \rangle = (2\pi)^3 \sum_{\mathbf{n}} \int \mathrm{d}^3 \mathbf{J} \, \omega_{\mathbf{n}}(\mathbf{J}) N_{\mathbf{n}}(\mathbf{J}), \tag{3.37}$$

where

$$N_{\mathbf{n}}(\mathbf{J}) = -\tfrac{1}{2}\left\{ \pi^2 (\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0)|\Phi_{\mathbf{n}}(\omega_{\mathbf{n}})|^2 + \frac{|g_{\mathbf{n}}(\omega_{\mathbf{n}})|^2}{\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0} \right\} \tag{3.38}$$

is the phase-space density of 'plasmons' associated with wavevector $\mathbf{n}$ in the terminology of Hamilton and Heinemann (2020). Remarkably, this expression for the plasmon density involves only the driving term $g_{\mathbf{n}}$ and the resulting spatial structure $\Phi_{\mathbf{n}}$ of the component van Kampen modes – there is no mention of the kinetic energy of the driven stars. At first sight this is odd because in a stable system disturbances fade while E is constant. The equation works because the potentials in question belong not to the disturbance but to its constituent van Kampen modes, which do not change but nevertheless cause the disturbance to fade as they become more and more evenly distributed in phase.

3.5 Discussion

There are close parallels between formulae derived here for general systems and ones presented in Chapter 2 for ergodic systems. The essential difference is the universal replacement of $\mathrm{d}f_0/\mathrm{d}H$ in Chapter 2 by $\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0 / \mathbf{n} \cdot \boldsymbol{\Omega}$ as here. The other differences are more superficial, being caused by formulae in Chapter 2 being derived from an operator that gives $\partial^2 f/\partial t^2$ rather than one that gives $\partial f/\partial t$. This change leads to $\delta\big((\mathbf{n} \cdot \boldsymbol{\Omega})^2 - \omega^2\big)$ in Chapter 2 being replaced by $\delta(\mathbf{n} \cdot \boldsymbol{\Omega} - \omega)$. These symbols have different dimensions:

$$\delta\big((\mathbf{n} \cdot \boldsymbol{\Omega})^2 - \omega^2\big) = \delta\big((\mathbf{n} \cdot \boldsymbol{\Omega} - \omega)(\mathbf{n} \cdot \boldsymbol{\Omega} + \omega)\big) \simeq \frac{\delta(\mathbf{n} \cdot \boldsymbol{\Omega} - \omega)}{2\omega}. \tag{3.39}$$

As a consequence, $g_{\mathbf{n}}/2\omega$ in Chapter 2 is equivalent to $g_{\mathbf{n}}$ here.

In Chapter 2 we explored the consequences of reducing the self-gravity of a system by reducing the masses of its particles by a factor $\xi < 1$ and introducing a fraction $(1 - \xi)$ of the mean-field potential. This operation modified the structure of the system's van Kampen modes by suppressing the dressing of the potential of resonant stars. The ensuing discussion applies equally to the general equilibria treated here.

In Chapter 2 we stressed that Landau 'modes' are not really modes but zeroes of the function $|\mathbf{M}(\omega)|$. When such a zero lies just below the real axis (a 'weakly damped mode') the van Kampen modes on the adjacent stretch of the real axis are heavily dressed and will make a large contribution to the system's evolution. The time required for these modes to drift out of phase, and the associated disturbance to fade in real space, is inversely proportional to extent of the heavily dressed section of the real line, and therefore inversely proportional to the distance of the zero from the real line. This interpretation of 'Landau damping' applies equally to the general systems discussed here.

Although Mathur (1990) did not display an explicit example, he demonstrated the logical possibility of marginally stable systems, that is ones that can oscillate for ever. This possibility arises when frequencies of the form $\mathbf{n} \cdot \mathbf{\Omega}(\mathbf{J})$ do not extend from a maximum frequency down to zero. The easiest way to engineer a gap in the frequency coverage is to impose a limit on the spatial extent of the system, so there is a minimum orbital frequency Ω_{min}, and to make the system effectively one-dimensional so small frequencies cannot be constructed by differencing frequencies greater than Ω_{min}. When these conditions are satisfied, Mathur (1990) shows that f_0 can be devised such that the matrix $\mathbf{M}$ of equation (3.32) has vanishing determinant in a gap. Consequently, non-zero $\mathbf{A}$ can then be found even though $\mathbf{B}$ vanishes because we are in a gap. In these exceptional circumstances, a complete set of modes comprises the van Kampen modes plus any stable/unstable pairs of modes and any marginally stable modes in gaps.

3.6 Conclusions

We have extended results presented in Chapter 2 from ergodic systems to systems with DFs of the form $f(\mathbf{J})$. This is a major extension because ergodic systems probably do not occur in nature while many galaxies and star clusters will have DFs that can be closely approximated by $f(\mathbf{J})$.

We derived equation (3.23) for the energy of a disturbance to a general system, and motivated by this result defined a bilinear form on the space of DFs such that the energy of a disturbance is the norm of its DF.

Next we extended the concept of van Kampen modes to general systems. They exist for essentially the same range of real frequencies, and have the same physical interpretation and energy-additivity as in the ergodic case. There is again a sharp distinction between van Kampen modes, which have frequencies that lie in a real continuum, and classical modes, which have isolated, and generally complex, frequencies. When a system is time-reversible (lacks rotation) and stable, it can only have van Kampen modes.

The mathematical apparatus used here is simpler and less powerful than that deployed in Chapter 2. On the plus side, the formulae presented here are somewhat simpler than the corresponding formulae in Chapter 2. The downsides are (i) that modes do not emerge as eigenfunctions of a Hermitian operator; (ii) we have not shown that all van Kampen modes have positive energy. The lack of a connection to a Hermitian operator deprives us of the ability both to argue for the completeness of the modes and to confine the normal-mode frequencies to the axes of the complex plane as we could in the case of ergodic systems. Nonetheless, at least in the case of systems with $f_0(\mathbf{J})$ that are unchanged by reversing all velocities, it seems likely that the frequencies are confined in the same way. This topic will be a major theme of in the fourth contribution (not included in this thesis) in this series.

The existence of modes with negative energy would not have important consequences on the dynamical timescale, because modes evolve independently of one another so long as the linear approximation holds. On the longer

timescales associated with terms quadratic in the disturbance (the 'two-body' timescale), a system is likely to be rendered secularly unstable by the existence of negative-energy modes because non-linear terms could transfer energy from negative-energy modes to positive-energy terms and thus causing the amplitudes of both types of mode to increase. We have only shown that we cannot exclude the possibility of these modes existing: they could very simply not.

3.7 Appendix: The bilinear form evaluated on van Kampen modes

Here we prove that the bilinear form (3.27) evaluated on two van Kampen modes can be expressed in the form (3.35). The product is the sum $K + \frac{1}{2}P'$ of kinetic- and potential-energy terms, where[2]

$$K[f, \widetilde{f}] \equiv -\frac{(2\pi)^3}{2} \sum_{\mathbf{n}} \int \mathrm{d}^3 \mathbf{J} \, \frac{\mathbf{n} \cdot \mathbf{\Omega}}{\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0} f_{\mathbf{n}}^* \widetilde{f}_{\mathbf{n}}$$

$$P'[f, \widetilde{f}] \equiv -\mathcal{E} \sum_{\alpha} A_{\alpha}^* \widetilde{A}_{\alpha}. \tag{3.40}$$

Using equation (3.30) for the DF of a mode, we have

$$K = -\frac{(2\pi)^3}{2} \sum_{\mathbf{n}} \int \mathrm{d}^3 \mathbf{J} \frac{\mathbf{n} \cdot \mathbf{\Omega}}{\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0}$$
$$\times \left\{ \mathcal{P}\left(\frac{\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0}{\mathbf{n} \cdot \mathbf{\Omega} - \omega} \right) \Phi_{\mathbf{n}}^* + g_{\mathbf{n}}^* \delta(\mathbf{n} \cdot \mathbf{\Omega} - \omega) \right\}$$
$$\times \left\{ \mathcal{P}\left(\frac{\mathbf{n} \cdot \nabla_{\mathbf{J}} f_0}{\mathbf{n} \cdot \mathbf{\Omega} - \widetilde{\omega}} \right) \widetilde{\Phi}_{\mathbf{n}} + \widetilde{g}_{\mathbf{n}} \delta(\mathbf{n} \cdot \mathbf{\Omega} - \widetilde{\omega}). \right\} \tag{3.41}$$

---

[2] P' is the potential energy of the system defined by f in the potential generated by $\widetilde{f}$ or vice versa [eqn. 3.10]. The self-energy $P = \frac{1}{2}P'[f, f]$.

With identities (3.47) and (3.48) the product of the two principal values yields

$$\int d^3\mathbf{J}\, \frac{\mathbf{n}\cdot\mathbf{\Omega}}{\mathbf{n}\cdot\nabla_\mathbf{J} f_0}\mathcal{P}(.)\mathcal{P}(.) = \int d^3\mathbf{J}\, \mathbf{n}\cdot\nabla_\mathbf{J} f_0 \Phi_\mathbf{n}^* \widetilde{\Phi}_\mathbf{n}$$
$$\times \left[\mathcal{P}\Big(\frac{1}{\omega-\widetilde{\omega}}\Big)\left\{\mathcal{P}\Big(\frac{\omega}{\mathbf{n}\cdot\mathbf{\Omega}-\omega}\Big) - \mathcal{P}\Big(\frac{\widetilde{\omega}}{\mathbf{n}\cdot\mathbf{\Omega}-\widetilde{\omega}}\Big)\right\} \right.$$
$$\left. + \pi^2\omega\delta(\mathbf{n}\cdot\mathbf{\Omega}-\omega)\delta(\omega-\widetilde{\omega}) \right]. \tag{3.42}$$

The cross terms yield

$$\int d^3\mathbf{J}\, \left\{ \mathcal{P}\Big(\frac{\mathbf{n}\cdot\mathbf{\Omega}}{\mathbf{n}\cdot\mathbf{\Omega}-\omega}\Big)\Phi_\mathbf{n}^*\widetilde{g}_\mathbf{n}\delta(\mathbf{n}\cdot\mathbf{\Omega}-\widetilde{\omega}) \right.$$
$$\left. + \mathcal{P}\Big(\frac{\mathbf{n}\cdot\mathbf{\Omega}}{\mathbf{n}\cdot\mathbf{\Omega}-\widetilde{\omega}}\Big)\widetilde{\Phi}_\mathbf{n} g_\mathbf{n}^*\delta(\mathbf{n}\cdot\mathbf{\Omega}-\omega). \right\} \tag{3.43}$$

The identity (3.49) enables us to rewrite this in the form

$$\int d^3\mathbf{J}\, \mathcal{P}\Big(\frac{1}{\omega-\widetilde{\omega}}\Big) \int d^3\mathbf{J}\, \left\{ \widetilde{\omega}\Phi_\mathbf{n}^*\widetilde{g}_\mathbf{n}\delta(\mathbf{n}\cdot\mathbf{\Omega}-\widetilde{\omega}) \right.$$
$$\left. - \omega\widetilde{\Phi}_\mathbf{n} g_\mathbf{n}^*\delta(\mathbf{n}\cdot\mathbf{\Omega}-\omega) \right\}. \tag{3.44}$$

When the cross terms are added to the reduced product of principal values (3.42), the coefficient of $\mathcal{P}\big(1/(\omega-\widetilde{\omega})\big)$ is

$$\left\{\mathcal{P}\Big(\frac{\mathbf{n}\cdot\nabla_\mathbf{J} f_0}{\mathbf{n}\cdot\mathbf{\Omega}-\omega}\Big)\Phi_\mathbf{n}^* + g_n^*\delta(\mathbf{n}\cdot\mathbf{\Omega}-\omega)\right\}\omega\widetilde{\Phi}_\mathbf{n}$$
$$-\left\{\mathcal{P}\Big(\frac{\mathbf{n}\cdot\nabla_\mathbf{J} f_0}{\mathbf{n}\cdot\mathbf{\Omega}-\widetilde{\omega}}\Big)\widetilde{\Phi}_\mathbf{n} + \widetilde{g}_n\delta(\mathbf{n}\cdot\mathbf{\Omega}-\widetilde{\omega})\right\}\widetilde{\omega}\Phi_\mathbf{n}^*$$
$$= f_\mathbf{n}^*\omega\widetilde{\Phi}_\mathbf{n} - \widetilde{f}_\mathbf{n}\widetilde{\omega}\Phi_\mathbf{n}^*. \tag{3.45}$$

When this expression is integrated wrt $\mathbf{J}$, summed over $\mathbf{n}$, and multiplied by $(2\pi)^3$, we obtain $\omega-\omega'$ times the mutual potential energy P' (eqn. 3.10). Hence the denominator of the principal value symbol in equation (3.44) is cancelled. Gathering together the contributions to K, we have this potential-energy term, the product of δ-functions in equation (3.42) and a similar term arising from

the product of $g_{\mathbf{n}}^*$ and $\widetilde{g}_{\mathbf{n}}$ in equation (3.41). Thus

$$K = -\tfrac{1}{2}P' - \frac{(2\pi)^3}{2}\sum_{\mathbf{n}}\int \mathrm{d}^3\mathbf{J}\,\omega\left(\pi^2\mathbf{n}\cdot\nabla_{\mathbf{J}}f_0\Phi_{\mathbf{n}}^*\widetilde{\Phi}_{\mathbf{n}}\right.$$

$$\left.+\frac{\mathbf{n}\cdot\mathbf{\Omega}}{\mathbf{n}\cdot\nabla_{\mathbf{J}}f_0}g_{\mathbf{n}}^*\widetilde{g}_{\mathbf{n}}\right)\delta(\mathbf{n}\cdot\mathbf{\Omega}-\omega)\delta(\omega-\widetilde{\omega}). \tag{3.46}$$

Thus $\langle f|\widetilde{f}\rangle = K + \tfrac{1}{2}P'$ is indeed given equation (3.35).

3.8 Appendix: Identities involving principal values

Here we list three identities that are required in Appendix 3.7. For proofs see (e.g. Ramos and White 2018).

$$\mathcal{P}\left(\frac{1}{x-x_1}\right)\mathcal{P}\left(\frac{1}{x-x_2}\right) =$$
$$\mathcal{P}\left(\frac{1}{x_1-x_2}\right)\left\{\mathcal{P}\left(\frac{1}{x-x_1}\right)-\mathcal{P}\left(\frac{1}{x-x_2}\right)\right\}$$
$$+\pi^2\delta(x-x_1)\delta(x_1-x_2), \tag{3.47}$$

$$\mathcal{P}\left(\frac{x}{x-x_1}\right) = 1 + \mathcal{P}\left(\frac{x_1}{x-x_1}\right). \tag{3.48}$$

$$\mathcal{P}\left(\frac{A}{x-x_2}\right)\delta(x-x_1)+\mathcal{P}\left(\frac{B}{x-x_1}\right)\delta(x-x_2)$$
$$= \mathcal{P}\left(\frac{1}{x_1-x_2}\right)\left\{A\delta(x-x_1)-B\delta(x-x_2)\right\}. \tag{3.49}$$

Chapter 4

Probabilistic Distribution Functions[1]

4.1   Introduction

Galaxies and star clusters are in approximate states of equilibrium and have for decades been fitted to models in which the distribution function (DF) $f(\mathbf{x}, \mathbf{v})$ of their constituent particles (stars, dark-matter particles) are steady-state solutions of the collisionless Boltzmann equation (CBE). The advent of massive simulations of galaxy formation (e.g. Laporte et al. 2019) and detailed data from the Gaia mission (Gaia Collaboration and Brown 2018) and large integral field units such as MUSE (e.g. Vitral and Mamon 2021) have stimulated interest in non-equilibrium features of galaxies, especially the Milky Way (e.g. Antoja et al. 2018).

It is likely that equilibrium models will not suffice to explain the exquisite data that Gaia is delivering because the data, effectively taken at one instant of a cluster's life, capture a specific fluctuation. Hence we now need models that embrace fluctuations. Clearly this can only be done in a statistical sense: the requirement is to predict which deviations from a mean-field model are likely. That is, we need to assign probabilities to distribution functions (Magorrian 2006), and the theory will be tested by comparing with observations its pre-

---

[1]This chapter is taken from Lau and Binney (2021d).

dictions for populations of clusters or galaxies.

The dynamics of a crystal are often best studied by going to the continuum limit in which quantities like the displacement $\boldsymbol{\xi}$ of atoms from equilibrium positions and the orientations $\mathbf{s}$ of spins are defined at all $\mathbf{x}$. Then the state of the system is defined by fields such as $\boldsymbol{\xi}(\mathbf{x})$ or $\mathbf{s}(\mathbf{x})$. Similarly, in stellar dynamics we want to characterise the disturbance of a cluster by the field $f(\mathbf{w})$, where $\mathbf{w} = (\mathbf{x}, \mathbf{v})$ is a location in phase space. Then we want to compute expectation values of observables $\mathcal{O}$ by taking expectation values over all possible fields f: doing so we are computing a double expectation: first $\langle \mathcal{O} \rangle_f = \int \mathrm{d}^6 \mathbf{w}\, f(\mathbf{w}) \mathcal{O}(\mathbf{w})$ and then an average of these averages weighted by the probability of each DF f.

An example of an important observable to predict in this way is the variance of a suitable dipole moment to quantify the tendency for the densest part of a cluster to be offset from the cluster's barycentre. In this case one would take $\mathcal{O}$ to be the product of a function $R(r)$ that has opposite signs at $r = 0$ and $r \to \infty$ and an $\ell = 1$ spherical harmonic:

$$\mathcal{O}(\mathbf{w}) = R(r) Y_1^m(\theta, \phi). \tag{4.1}$$

In the case of an isolated, non-rotating cluster $\langle \mathcal{O} \rangle_f$ would then be a random variable with vanishing mean and a dispersion that would be independent of the azimuthal quantum number m. The variance of $\langle \mathcal{O} \rangle_f$ would be a prediction that could be tested from observations of a sufficient number of real or simulated clusters (e.g. Lau and Binney 2019, Heggie et al. 2020).

This programme requires a rule for assigning a priori probabilities to possible fields $f(\mathbf{w})$. In classical statistical mechanics the analogous rule is inferred by noting that the a priori probability of some range of phase-space locations $\mathbf{w}$ must be independent of time as the system evolves undisturbed, and Liouville's theorem ensures that this condition is satisfied if a priori probability is proportional to the volume element $\mathrm{d}^6 \mathbf{w} = \mathrm{d}^3 \mathbf{q}\, \mathrm{d}^3 \mathbf{p}$ defined by a system of canonical coordinates $(\mathbf{q}, \mathbf{p})$ – here it is important that (1) any system of canonical coordinates assigns probability in the same way, and (2) the assignment

is time-independent because Hamilton's equations effect a series of canonical transformations.

Guided by this analogy, we seek a system of canonical coordinates for the space of possible DFs. Canonical coordinates are defined to be those in which Poisson brackets $[f, g]$ have the canonical form

$$[f, g] = \sum_i \left(\frac{\partial f}{\partial \mathbf{q}} \cdot \frac{\partial g}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}} \cdot \frac{\partial g}{\partial \mathbf{q}} \right). \tag{4.2}$$

Poisson brackets impose a symplectic structure on phase space and canonical coordinates are privileged coordinates with respect to this structure in the same way that Cartesian coordinates are privileged with respect to a Euclidean metric (e.g. Arnold 1989). So the key to defining a field theory for stellar dynamics is imposing a symplectic structure on the space of all possible DFs. That is, we must define an antisymmetric bilinear form $\{,\}$ that yields a functional on the space of DFs when the two slots are filled by two DFs. The standard phase-space Poisson bracket does not fulfil this role because $[f, g]$ is a function on phase space rather than a functional. In Section 4.3 we identify a symplectic structure for the space of DFs. This structure then allows us to identify in Section 4.3.2 canonical coordinates for the space of DFs.

Chapters 2 and 3 argue that the fluctuations of a stellar system are best tackled in terms of the system's van Kampen modes. They showed that a system's excitation energy is the sum of the energies invested in each of its van Kampen modes. The energy of a growing or decaying mode is identically zero and in Section 4.4 we show that when the energy of a stable mode is expressed in terms of canonical coordinates, it takes the form of a sum of harmonic-oscillator Hamiltonians. This result suggests that it may be possible to apply to clusters the methods of classical equilibrium statistical mechanics.

4.2 Mathematical background and notation

Here we introduce two vital concepts and define our notation.

4.2.1 Angle-action variables

Actions J_i are constants of motion that can serve as canonical phase-space coordinates. Their conjugate variables, the angles θ_i, increase linearly in time, so $\boldsymbol{\theta}(t) = \boldsymbol{\theta}(0) + \boldsymbol{\Omega}t$. A particles' Hamiltonian $H(\mathbf{x}, \mathbf{v})$ is a function $H(\mathbf{J})$ of the actions only and the frequencies Ω_i that control the rates of increase of the angles are given by $\boldsymbol{\Omega} = \partial H/\partial \mathbf{J}$. Since angle-action variables are canonical coordinates, the element of phase-space volume $\mathrm{d}^6\mathbf{w} = \mathrm{d}^3\mathbf{x}\,\mathrm{d}^3\mathbf{v} = \mathrm{d}^3\boldsymbol{\theta}\,\mathrm{d}^3\mathbf{J}$. Functions on phase space can be expressed as Fourier series:

$$h(\mathbf{w}) = \sum_{\mathbf{k}} h_{\mathbf{k}}(\mathbf{J})\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\boldsymbol{\theta}} \; ; \; h_{\mathbf{k}}(\mathbf{J}) = \int \frac{\mathrm{d}^3\boldsymbol{\theta}}{(2\pi)^3}\mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\boldsymbol{\theta}}h(\mathbf{w}). \tag{4.3}$$

Note that for real h, $h_{-\mathbf{k}} = h_{\mathbf{k}}^*$.

4.2.2 van Kampen modes

The linearised CBE is time-reversal symmetric, so group-theoretic arguments suggest that it should have a complete set of solutions with time dependence $\mathrm{e}^{\mathrm{i}\omega t}$, with ω possibly complex. These solutions are called van Kampen modes. Chapter 2 shows that these solutions are the eigenfunctions of a Hermitian operator on phase space K, with eigenvalue ω^2. Since the eigenvalues of a Hermitian operator are all real, ω is either real or purely imaginary. Time-reversal symmetry implies that if there is a van Kampen mode with frequency $\omega = \mathrm{i}y$ with y real, then there is another mode with frequency $\omega^* = -\mathrm{i}y$. Hence modes are either purely oscillatory or exponentially growing/decaying in pairs. The real frequencies form a continuum, while the imaginary frequencies are isolated.

4.3 A symplectic structure for the space of possible DFs

In this section we identify the symplectic structure for the space of DFs. Coordinates for the space of DFs are numbers that characterise a DF in the same way that quantum amplitudes $a_i \equiv \langle E_i|\psi\rangle$ are numbers that characterise the

state $|\psi\rangle$ of a quantum system. Suitable numbers can be extracted by taking an inner product with a function $a_i(\mathbf{w})$:

$$\hat{a}_i[f] \equiv (a_i, f) \equiv \int \mathrm{d}^6\mathbf{w} \, a_i^*(\mathbf{w}) f(\mathbf{w}). \tag{4.4}$$

Here the hat in $\hat{a}$ implies a functional on the space of DFs, and the square brackets imply that the argument is a function on phase space rather than a number. One possible choice of function that defines $\hat{a}_i$ is $a_{\mathbf{w}'}(\mathbf{w}) = \delta(\mathbf{w} - \mathbf{w}')$. Then $\hat{a}_{\mathbf{w}'}[f] = f(\mathbf{w}')$ is just the value of the DF at the location $\mathbf{w}'$ associated with this coordinate. Another widely employed choice is $a_{\mathbf{kJ}'} = \delta(\mathbf{J}' - \mathbf{J})\mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\boldsymbol{\theta}}/(2\pi)^3$, which yields $\hat{a}_{\mathbf{kJ}'}[f] = f_{\mathbf{k}}(\mathbf{J}')$. We show below that via equation (4.21) the $f_{\mathbf{k}}$ yield canonical coordinates for the space of DFs.

We are interested in how coordinates vary as we move about the space of DFs, so we need the functional derivative

$$\frac{\delta \hat{a}_i}{\delta f(\mathbf{w}')} \equiv \int \mathrm{d}^6\mathbf{w} \, a_i^*(\mathbf{w}) \delta(\mathbf{w} - \mathbf{w}') = a_i^*(\mathbf{w}'). \tag{4.5}$$

Note that taking a functional derivative of a coordinate $\hat{a}_i$ (a functional) we obtain a function on phase space.

Inspired by Morrison (1980), we construct a symplectic operator such that CBE can be written in the form of Hamilton's equations, namely

$$\frac{\partial \hat{f}}{\partial t} = \{\hat{f}, \hat{H}_{\mathrm{s}}\}, \tag{4.6}$$

where $\hat{f}$ is a functional such as $\hat{a}_i$ that characterises the system's DF and[2]

$$\hat{H}_{\mathrm{s}}[f] \equiv \tfrac{1}{2} \int \mathrm{d}^6\mathbf{w} \, f(\mathbf{w})\big(v^2 + \Phi(\mathbf{w})\big), \tag{4.7}$$

with

$$\Phi(\mathbf{w}) = -GM \int \mathrm{d}^6\mathbf{w}' \, \frac{f(\mathbf{w}')}{|\mathbf{x} - \mathbf{x}'|}. \tag{4.8}$$

In equation (4.6) both sides are functionals and the Poisson bracket on the right is defined by

$$\{\hat{U}, \hat{V}\}[f] \equiv \int \mathrm{d}^6\mathbf{w} \, f(\mathbf{w}) \left[\frac{\delta \hat{U}}{\delta f(\mathbf{w})}, \frac{\delta \hat{V}}{\delta f(\mathbf{w})} \right], \tag{4.9}$$

where the square brackets in the integral denote a conventional Poisson bracket.

---

[2]Φ appears with a factor half in the Hamiltonian because it is the potential self-consistently generated by the DF rather than an externally imposed potential.

4.3.1 Equivalence of equation (4.6) and the CBE

The functional derivative of $\hat{H}_{\rm s}$ is

$$
\begin{aligned}
\frac{\delta \hat{H}_{\rm s}}{\delta f(\mathbf{w}')} &= \tfrac{1}{2}\left(v'^2 + \Phi(\mathbf{x}')\right) + \tfrac{1}{2}\int \mathrm{d}^6\mathbf{w}\, f(\mathbf{w})\frac{\delta \Phi(\mathbf{w})}{\delta f(\mathbf{w}')} \\
&= \tfrac{1}{2}\left(v'^2 + \Phi(\mathbf{x}')\right) - \tfrac{1}{2}GM\int \mathrm{d}^6\mathbf{w}\, \frac{f(\mathbf{w})}{|\mathbf{x}-\mathbf{x}'|} \\
&= \tfrac{1}{2}v'^2 + \Phi(\mathbf{x}'),
\end{aligned}
\tag{4.10}
$$

which is the Hamiltonian of a single star. When we substitute the Poisson bracket's definition (4.9) into equation (4.6) and use our expression for the functional derivative of $\hat{H}_{\rm s}$, we find

$$
\frac{\partial \hat{f}}{\partial t}[f] = \int \mathrm{d}^6\mathbf{w}\, f(\mathbf{w})\left[\frac{\delta \hat{f}}{\delta f(\mathbf{w})}, H_{\rm s}(\mathbf{w})\right],
\tag{4.11}
$$

This equation holds for any choice of functional (coordinate) $\hat{f}$ on DF space. If we set $\hat{f}$ equal to the functional $\hat{a}_{\mathbf{w}'}$ with $a_{\mathbf{w}'}(\mathbf{w}) = \delta(\mathbf{w}-\mathbf{w}')$ explored above, then the number $\hat{f}[f]$ becomes $f(\mathbf{w}')$ and the functional derivative $\delta\hat{f}/\delta f(\mathbf{w})$ becomes $\delta(\mathbf{w}-\mathbf{w}')$, so we have

$$
\frac{\partial f(\mathbf{w}')}{\partial t} = \int \mathrm{d}^6\mathbf{w}\, f(\mathbf{w})[\delta(\mathbf{w}-\mathbf{w}'), H(\mathbf{w})].
\tag{4.12}
$$

For any three functions we have that $\int \mathrm{d}^6\mathbf{w}\, f[g,h] = \int \mathrm{d}^6\mathbf{w}\, g[h,f]$ (provided one or more vanishes at infinity), so equation (4.12) yields the conventional CBE

$$
\frac{\partial f(\mathbf{w}')}{\partial t} = [H(\mathbf{w}'), f(\mathbf{w}')],
\tag{4.13}
$$

The opposing signs in these two forms of the CBE (4.6) and (4.13) reflect the difference between Hamilton's equation $\dot{q} = [q, H]$ and the conventional CBE $0 = \dot{f} = \partial_t f + [f, H] \Rightarrow \partial_t f = -[f, H]$. That is, equation (4.6) is the equation of motion of the DF's coordinates, while equation (4.13) gives the rate of change of the value of f at fixed $\mathbf{w}$.

Above we derived the conventional CBE (4.13) from equation (4.6), but reversing the chain or arguments one can also show that equation (4.6) follows

from the conventional CBE; the statements $\partial_t \hat{f} = \{\hat{f}, \hat{H}_s\}$ and $\partial_t f = -[f, H]$ are equivalent.

4.3.2 Application to equilibrium systems

We now examine the structure of the functional Poisson bracket when evaluated on a DF that differs from that of an equilibrium model only by virtue of a fluctuation. That is, we consider DFs of the form

$$f(\mathbf{w}) = f_0(\mathbf{J}) + f_1(\boldsymbol{\theta}, \mathbf{J}). \tag{4.14}$$

In this case the Poisson bracket (4.9) can be written

$$
\{\hat{U}, \hat{V}\}[f] = \int \mathrm{d}^6\mathbf{w} \, \frac{\delta\hat{U}}{\delta f(\mathbf{w})} \left[\frac{\delta\hat{V}}{\delta f(\mathbf{w})}, f \right]
$$
$$
= \int \mathrm{d}^6\mathbf{w} \, \frac{\delta\hat{U}}{\delta f(\mathbf{w})} \left(\frac{\partial}{\partial\boldsymbol{\theta}} \frac{\delta\hat{V}}{\delta f(\mathbf{w})} \cdot \frac{\partial f}{\partial\mathbf{J}} - \frac{\partial}{\partial\mathbf{J}} \frac{\delta\hat{V}}{\delta f(\mathbf{w})} \cdot \frac{\partial f}{\partial\boldsymbol{\theta}} \right). \tag{4.15}
$$

The second term in the round bracket is smaller than the first by $\mathrm{O}(f_1/f_0)$, so to leading order we have

$$
\{\hat{U}, \hat{V}\}[f] = \int \mathrm{d}^6\mathbf{w} \, \frac{\delta\hat{U}}{\delta f(\mathbf{w})} \frac{\partial f_0}{\partial\mathbf{J}} \cdot \frac{\partial}{\partial\boldsymbol{\theta}} \frac{\delta\hat{V}}{\delta f(\mathbf{w})} \tag{4.16}
$$

If we define

$$
x_\mathbf{k} = \Re(f_\mathbf{k}) = \int \frac{\mathrm{d}^3\boldsymbol{\theta}}{(2\pi)^3} f \cos(\mathbf{k}\cdot\boldsymbol{\theta})
$$
$$
y_\mathbf{k} = \Im(f_\mathbf{k}) = -\int \frac{\mathrm{d}^3\boldsymbol{\theta}}{(2\pi)^3} f \sin(\mathbf{k}\cdot\boldsymbol{\theta}), \tag{4.17}
$$

then we can restrict sums over $\mathbf{k}$ to a half-plane, such as $k_1 \geq 0$ and

$$
\frac{\delta x_\mathbf{k}(\mathbf{J})}{\delta f(\mathbf{w})} = \frac{\cos\mathbf{k}\cdot\boldsymbol{\theta}}{(2\pi)^3} \quad \text{and} \quad \frac{\delta y_\mathbf{k}(\mathbf{J})}{\delta f(\mathbf{w})} = -\frac{\sin\mathbf{k}\cdot\boldsymbol{\theta}}{(2\pi)^3}. \tag{4.18}
$$

Now applying the chain rule to (4.16) yields

$$
\{\hat{U},\hat{V}\}[f] = \frac{1}{(2\pi)^6} \int \mathrm{d}^6\mathbf{w}
$$

$$
\sum_{\mathbf{k}} \left(\frac{\delta\hat{U}}{\delta x_{\mathbf{k}}(\mathbf{J})} \cos(\mathbf{k}\cdot\boldsymbol{\theta}) - \frac{\delta\hat{U}}{\delta y_{\mathbf{k}}(\mathbf{J})} \sin(\mathbf{k}\cdot\boldsymbol{\theta}) \right) \frac{\partial f_0}{\partial \mathbf{J}} \cdot \frac{\partial}{\partial\boldsymbol{\theta}}
$$

$$
\sum_{\mathbf{k}'} \left(\frac{\delta\hat{V}}{\delta x_{\mathbf{k}'}(\mathbf{J})} \cos(\mathbf{k}'\cdot\boldsymbol{\theta}) - \frac{\delta\hat{V}}{\delta y_{\mathbf{k}'}(\mathbf{J})} \sin(\mathbf{k}'\cdot\boldsymbol{\theta}) \right)
$$

$$
= \frac{1}{2(2\pi)^3} \int \mathrm{d}^3\mathbf{J} \sum_{\mathbf{k}} \mathbf{k}\cdot\frac{\partial f_0}{\partial\mathbf{J}} \left(-\frac{\delta\hat{U}}{\delta x_{\mathbf{k}}}\frac{\delta\hat{V}}{\delta y_{\mathbf{k}}} + \frac{\delta\hat{U}}{\delta x_{\mathbf{k}}}\frac{\delta\hat{V}}{\delta y_{\mathbf{k}}} \right).
$$

$$
(4.19)
$$

To bring the Poisson bracket (4.9) to the canonical form

$$
\{\hat{U},\hat{V}\}[f] = \int \mathrm{d}^6\mathbf{w} \sum_{\mathbf{k}} \left(\frac{\delta\hat{U}}{\delta q_{\mathbf{k}}}\frac{\delta\hat{V}}{\delta p_{\mathbf{k}}} - \frac{\delta\hat{U}}{\delta p_{\mathbf{k}}}\frac{\delta\hat{V}}{\delta q_{\mathbf{k}}} \right) \tag{4.20}
$$

we introduce new coordinates

$$
q_{\mathbf{k}} \equiv u_{\mathbf{k}} x_{\mathbf{k}} \quad ; \quad p_{\mathbf{k}} \equiv u_{\mathbf{k}} y_{\mathbf{k}}, \tag{4.21}
$$

where

$$
u_{\mathbf{k}} \equiv \sqrt{\frac{2(2\pi)^6}{-\mathbf{k}\cdot\partial f_0/\partial\mathbf{J}}}. \tag{4.22}
$$

and restrict the sum over $\mathbf{k}$ to the half of k space in which $\mathbf{k}\cdot\partial f_0/\partial\mathbf{J} < 0$.

To summarise: the Fourier coefficients $q_{\mathbf{k}}$ and $p_{\mathbf{k}}$ of DFs have emerged as canonical coordinates for the space of DFs. If we assign a priori probability to sets of DFs according to their volume elements $\int \mathrm{d}^3\mathbf{J} \prod_{\mathbf{k}} \mathrm{d}q_{\mathbf{k}}\mathrm{d}p_{\mathbf{k}}$, our assignment will be invariant under canonical transformation to new coordinates for DFs, and, crucially, the a priori probability of a set of DFs will remain unchanged as the DFs evolve according to the CBE. We have shown that $(\mathbf{q_k}, \mathbf{p_k})$ are canonical coordinates only to linear order, but the invariance of probabilities assigned by canonical coordinates is exact because the full CBE is of Hamiltonian form.

4.4 Ergodic systems

Here we derive some results for stable ergodic systems – equilibria with DFs that are functions of the single-particle Hamiltonian $H(\mathbf{J})$, so $f_0(H(\mathbf{J}))$. Antonov (1961) showed that to be stable the DF must satisfy $f_0'(H) < 0$.

4.4.1 Inner product

For such systems the natural inner product of the space of DFs is (e.g. Chapter 2)

$$\langle f|g \rangle \equiv \int \frac{\mathrm{d}^6 \mathbf{w}}{|f_0'(H)|} f^*(\mathbf{w}) g(\mathbf{w}). \tag{4.23}$$

It is simple to show that expressed in terms of Fourier components the product takes the form

$$\langle f|g \rangle = (2\pi)^3 \int \frac{\mathrm{d}^3 \mathbf{J}}{|f_0'|} \sum_{\mathbf{k}} f_{\mathbf{k}}^* g_{\mathbf{k}}. \tag{4.24}$$

4.4.2 Even and odd DFs

Antonov (1961) showed that it is useful to split the DFs of perturbed ergodic systems into parts even and odd in $\mathbf{v}$:

$$f_{\pm}(\mathbf{x}, \mathbf{v}) \equiv \tfrac{1}{2}\big\{ f(\mathbf{x}, \mathbf{v}) \pm f(\mathbf{x}, -\mathbf{v}) \big\}. \tag{4.25}$$

The unperturbed DF lies entirely within f_+ while f_- is entirely due to the perturbation. It is easy to show from the CBE that to first order in the perturbation $f_{\pm}$ are related by

$$\partial_t f_+ = -\mathbf{\Omega} \cdot \partial_{\boldsymbol{\theta}} f_-. \tag{4.26}$$

In the case of a normal mode

$$f(\mathbf{w}, t) = \sum_{\mathbf{k} \cdot \mathbf{\Omega} > 0} \left(f_{\mathbf{k}} \mathrm{e}^{\mathrm{i}(\mathbf{k} \cdot \boldsymbol{\theta} - \omega t)} + f_{\mathbf{k}}^* \mathrm{e}^{-\mathrm{i}(\mathbf{k} \cdot \boldsymbol{\theta} - \omega t)} \right), \tag{4.27}$$

where the sum is restricted to half of $\mathbf{k}$ space to compensate for the explicit inclusion of the complex conjugate of each term. When we substitute this expansion in equation (4.26) and equate coefficients of exponentials, we find

$$\omega f_{\mathbf{k}+} = \mathbf{k} \cdot \mathbf{\Omega} f_{\mathbf{k}-} \text{ and } \omega f_{\mathbf{k}+}^* = \mathbf{k} \cdot \mathbf{\Omega} f_{\mathbf{k}-}^* \ (\mathbf{k} \cdot \mathbf{\Omega} > 0). \tag{4.28}$$

Given that $f_{\mathbf{k}}^* = f_{-\mathbf{k}}$, a change of variable to $\mathbf{k}' \equiv -\mathbf{k}$ in the second relation yields

$$\omega f_{\mathbf{k}'+} = -\mathbf{k}' \cdot \mathbf{\Omega} f_{\mathbf{k}'-} \quad (\mathbf{k}' \cdot \mathbf{\Omega} < 0) \tag{4.29}$$

so we always have $\omega f_{\mathbf{k}+} = |\mathbf{k} \cdot \mathbf{\Omega}| f_{\mathbf{k}-}$. Hence,

$$f_{\mathbf{k}} = f_{\mathbf{k}+} + f_{\mathbf{k}-} = \left(\frac{|\mathbf{k} \cdot \mathbf{\Omega}|}{\omega} + 1 \right) f_{\mathbf{k}-}. \tag{4.30}$$

4.4.3 Energy of a mode

The inner product has the dimensions of energy, and indeed Chapter 2 shows that the energy of a van Kampen mode is

$$E_{\mathrm{mode}} = \langle f_- | f_- \rangle. \tag{4.31}$$

Crucially, the energies of van Kampen modes are additive because their odd DFs f_- are mutually orthogonal.

Equations (4.24) and (4.30) allow us to express E in terms of the Fourier components of the complete perturbed DF as

$$E_{\mathrm{mode}} = (2\pi)^3 \int \frac{\mathrm{d}^3 \mathbf{J}}{|f_0'|} \sum_{\mathbf{k}} \frac{|f_{\mathbf{k}}|^2}{(1 + |\mathbf{k} \cdot \mathbf{\Omega}|/\omega)^2}. \tag{4.32}$$

Rewritten in terms of the canonical variables (4.21) the energy is

$$E_{\mathrm{mode}} = \tfrac{1}{2} \int \mathrm{d}^3 \mathbf{J} \sum_{\mathbf{k} \cdot \mathbf{\Omega} > 0} \frac{\mathbf{k} \cdot \mathbf{\Omega}}{(1 + |\mathbf{k} \cdot \mathbf{\Omega}|/\omega)^2} \left(|q_{\mathbf{k}}|^2 + |p_{\mathbf{k}}|^2 \right). \tag{4.33}$$

E_{mode} has the form of the sum of the Hamiltonians of harmonic oscillators with frequencies $\mathbf{k} \cdot \mathbf{\Omega}/(1 + |\mathbf{k} \cdot \mathbf{\Omega}|/\omega)^2$. The energy of the entire system, being a sum of the energies of individual modes, is also a sum of harmonic-oscillator energies

$$\begin{aligned} E_{\mathrm{total}} &= \sum_{\mathrm{modes}} E_{\mathrm{mode}} \\ &= \tfrac{1}{2} \int \mathrm{d}^3 \mathbf{J} \sum_{\mathbf{k} \cdot \mathbf{\Omega} > 0} \sum_{\mathrm{modes}} \frac{\mathbf{k} \cdot \mathbf{\Omega}}{(1 + |\mathbf{k} \cdot \mathbf{\Omega}|/\omega)^2} \left(|q_{\mathbf{k}}|^2 + |p_{\mathbf{k}}|^2 \right). \end{aligned}$$

$$\tag{4.34}$$

The key to computing thermodynamic potentials is to express the system's Hamiltonian as a sum of harmonic-oscillator Hamiltonians because with that done the partition function (or the entropy $-k_\mathrm{B} \ln \Omega$) is readily evaluated. Equation (4.33) brings us closer to that goal but does not achieve it: the energy of a single mode is given in terms of contributions from many oscillators: each pair $(\mathbf{k}, \mathbf{J})$ corresponds to a different oscillator. Chapter 2 gives an expression for E_mode that involves an integral over just resonant tori, i.e., ones satisfying $\mathbf{k} \cdot \mathbf{\Omega} = \omega$, but this integral involves the mode's potential $\Phi[f]$, which is not easily computed. What is needed is a canonical transformation from $(\mathbf{q_k}[f], \mathbf{p_k}[f])$ to new functionals $(Q_i[f], P_i[f])$ such that the energy of the ith mode is $\Omega_i(Q_i^2 + P_i^2)$. Since the Hamiltonian is a quadratic function in both systems, the sought-after transformation could be linear. The new functionals $(Q_i[f], P_i[f])$ will encode the amplitude and phase of the contribution from the ith van Kampen mode that is required to build up the an arbitrary DF.

Morrison and Shadwick (1994) identified the required functionals in the case of a one-dimensional, homogeneous, electrostatic plasma, but their treatment does not immediately generalise to multiple spatial dimensions even in the case of a plasma. Perhaps a similar transform could be found that works in the three-dimensional case. If this could be done for an electrostatic plasma, it could almost certainly be adapted to the very similar case of the gravitating periodic cube (Barnes et al. 1986), which also offers scope for some interesting numerical experiments.

4.5   Conclusions

Galaxies and star clusters are in states that differ from the steady-state solutions to the CBE that have traditionally been used to interpret data. Differences between actual and idealised states can now be detected in the most precise modern data, so we need to extend stellar dynamics so these deviations are appropriately predicted. Predictions will generally be of a statistical nature: particular deviations will be assigned probabilities. The natural way

to do this is to assign probabilities to individual DFs. This needs to be done in a consistent manner. In particular, probabilities should not change when DFs evolve under the CBE. We have shown how this requirement can be met by identifying a symplectic structure and associated canonical coordinates for the space of DFs.

The ideal canonical coordinates for DF space would be functionals that determine the amplitude and phase of each van Kampen mode that is required to build up a given DF. These would be the angle-action cordinates of DF space. We have not identified these functionals, but we have identified one system of canonical coordinates, from which the angle-action coordinates might be derived through a canonical transformation. We have shown, moreover, that the energy of van Kampen mode is a quadratic function of the identified coordinates. This amounts to a significant step on the road to a theory of the thermodynamics of stellar systems.

Chapter 5

Boltzmann Equation Field Theory I: Ensemble Averages[1]

5.1   Introduction

The long-standing study of the dynamics of observed stellar systems began with E. C. Pickering (1891) who put forth the claim that different globular clusters (GCs) had similar "laws of distributions", referring to the distribution of stars inferred from 2D projected images obtained from the Draper Catalogue.

Plummer (1911) twenty years later proposed to "pursue the search for a physical basis on which the distribution of stars in clusters may be established". Under the presumption that spherical nebulae that are in convective equilibrium produce spherical globular clusters of the same mass distribution, he proposed a physically meaningful and demonstrably suitable fit (within $\pm 5\%$) to Pickering's data; his Plummer model.

Soon after, Eddington (1916) utilised Abel's theorem to translate 2D projected observations of GCs into 3D position-space, a result used by Plummer to produce his formulae connecting $f(\mathbf{x}, \mathbf{v}, t) = f(\mathbf{w}, t)$ the 6D distribution of stars in position and velocity space (henceforth phase space) to their mass distributions in position-space.

---

[1]This chapter is taken from Lau (2023).

And from thereon, the general methodology of fitting $\{\mathbf{w}_i\}$, the observed phase-space coordinates of a set of stars which are believed to form a gravitationally bound system with f, the distribution function (DF) of these stars, is as follows: First, intuit the symmetries of this observed system: e.g. the stars in the Milky Way lie (approximately) in a plane; thus a valid approximation would be to model galaxies as 2D thin disk. Secondly, compute a distribution function which obeys these symmetries. Finally, fit this symmetric distribution function to the observed data.

Examples of this methodology are found in nearly every system: Galaxies were first fit by the axisymmetric Mestel (1963) disc profile, which bellied the flat rotation curve found observationally and the Hernquist (1990) profile approximates De Vacouleur's law for the surface brightness of elliptical galaxies. In the realm of globular clusters, Henon (1959)'s isochrone distribution is the most general distribution of stars for which the radial period of orbits is purely a function of their energy. King (1966) championed the usage of a series of lowered isothermal spheres; systems with finite mass and extent and a simple truncated exponential dependence of the density of stars on the energy. Osipkov (1979)-Merritt (1985) distribution functions describe a simple way to introduce velocity-space anisotropy to initially isotropic, spherical systems.

These distribution functions are also known as mean-fields, $f(\mathbf{w}, t) = f_0(\mathbf{w})$, and share several commonalities:

Firstly, they are static in time. This means that states f_0 are assumed to be in equilibrium. Whether this is because equilibrium f_0 are simpler to calculate, or because of the long-standing hypothesis that systems must be stable to be observed in the present day is up to contention. Just looking at the extensive literature regarding streams of stars being tidally torn from globular clusters/dwarf galaxies by the Milky Way, density wave theory in the context of the development of spirals (see Lin and Shu (1964) who studied quasi-stationary spiral waves and Sellwood and Carlberg (2014) for a more recent contribution modelling transient spirals)/slowing bar theory (Chiba et al. 2020)

is that the assumption that observed systems are well fit by steady-state DF serves as a first approximation at best.

Secondly, they respect some spatial geometric symmetry. This approximation was doubtful even back in 1911, as Plummer wrote in reference to the fit of his model to an observed GC, 'it is clear that the counts are at least as well represented by the formula we have chosen as they are consistent with the fundamental hypothesis that we are dealing with a truly spherical distribution'. The natural asymmetry of astrophysical systems has been known of for over a century observationally and theoretically. Gravity is an attractive long-ranged force. Gravitationally stable systems generally exhibit system-scale fluctuations; observationally we see spiral arms, galactic bars, globular clusters with well developed dipolar asymmetries, fluctuations which are not captured by the mean-field symmetries.

I think it is clear that fitting f_0 to a set of observed particles $\{\mathbf{w}_i\}$ is not the right approach. To put it very simply, it should be clear that we do not know what the DF is! We know that there are very many possible distribution functions which may be fit to any set of observed particles, and we know we cannot claim certainty in knowing that any single f fits $\{\mathbf{w}_i\}$. So instead of choosing $f = f_0$ based on a physical argument (e.g. maximum entropy, or any of the other reasons conferred above), we know should be assigning probabilities $P[f]$ to f, based on those physical arguments. This insight—that it is better to accomodate our ignorance of the DF than to just choose what we think is sensible—is invoked earlier by Magorrian (2006) who applied it in criticism of maximum likelihood methods. He showed that the mass of supermassive black holes in toy galaxies for which mock observations have been obtained and trial DFs have been inferred via Bayesian inference, are better constrained if one selects for the mass which fits as many viable trial DFs as possible, and not for the mass which fits the most viable trial DF possible.

Now it must be noted that Pickering and Plummer conceived of their theories with only a handful of observations to draw from. Even until the

late 1950s, astrophysicists only had several tens of thousands of stars to draw measurements from. Their data was sufficiently coarse to be able to neglect variations in f, having neither the resolution nor the statistics to make claims regarding the substructures of astronomical objects, so they did not need to define such a $P[f]$. It is because we know more that a theory which assigns probabilities to different variations of f is necessary.

Gaia Collaboration et al. (2022), has revealed a dense network of asymmetries and large-scale substructure within and without our Galaxy; the Antoja et al. (2018) spiral, the Gaia-Enceladus sausage, or the newly coined 'X-shaped, Peanut, Boxy bar'. The corotation resonances of the bar severing the Hercules stream (see (Monari et al. 2019)), or the evidence for the slowing bar as stars trapped within bar resonances (Chiba and Schönrich 2021). The question of whether fluctuations matter, not just to the secular evolution of systems but also to what we see today was pertinent before GAIA, but is crucial to understand now.

So we need a variational theory in f, which allows us to predict the properties of fluctuations beyond the steady-state distributions: this is a field theory. In this chapter, I will derive such a theory which accounts for finiteness noise, gravitational correlations, and more, all under the framework of the CBE!

Section 5.2 describes essential mathematical tools to understand this chapter. Section 5.3 describes how fluctuations in f can be sourced from discreteness noise via an entropy argument. Section 5.4 describes how we can introduce desired structures into this noise. Section 5.5 describes how we can shape fluctuations into a perturbative field theory. Section 5.6 describes how we can reclaim observables using the field theory, Section 5.7 describes fluctuations corresponding to a system with a mean energy constraint and Section 5.8 shows how we can make predictions. Finally, Sections 5.9 and 5.10 summarise and deliberate over the predictions of this field theory.

5.2 Mathematical background

Here I introduce essential mathematical tools and establish my notation.

5.2.1 Poisson Brackets

The Poisson brackets can be computed as

$$[f, g] = \sum_i \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right). \tag{5.1}$$

where the pair of canonical phase-space coordinates $(\mathbf{q}, \mathbf{p})$ obey,

$$[p_i, p_j] = 0, \ [q_i, q_j] = 0, \ [q_i, p_j] = \delta_{ij} \tag{5.2}$$

In astrophysics, there are two important choices of such coordinates, the 3D Cartesian positions/velocities, $\mathbf{w} = (\mathbf{x}, \mathbf{v})$, and their generalisations, the angle/actions, $\mathbf{w} = (\boldsymbol{\theta}, \mathbf{J})$. The angle-action coordinates generalise positions and velocities for inhomogeneous systems. The actions $\mathbf{J}$ are constants of motion for orbits defined by the flow of the globally integrable Hamiltonian $H = H(\mathbf{J})$, whereas the angles $\boldsymbol{\theta}$ are the conjugates of $\mathbf{J}$, and thus obey $\boldsymbol{\theta}(t) = \boldsymbol{\theta}_0 + \boldsymbol{\Omega}t$, $\boldsymbol{\Omega} = \partial H / \partial \mathbf{J}$ by Hamilton's equation. Due to their canonicity, the phase-space element may be expressed $\mathrm{d}^6\mathbf{w} = \mathrm{d}^3\mathbf{x}\mathrm{d}^3\mathbf{v} = \mathrm{d}^3\boldsymbol{\theta}\mathrm{d}^3\mathbf{J}$.

While action-angle coordinates provide xa simple way to track how a star moves along its trajectory, their usage presumes that the Hamiltonian is globally integrable. A field theory integrates over all distribution functions, and general Hamiltonians only admit local integrability, (i.e. different parts of a system have different conserved quantities). Thus, I choose to use $\mathbf{w} = (\mathbf{x}, \mathbf{v})$ unless otherwise stated.

5.2.2 CBE

The collisionless Boltzmann equation governs the evolution of the one-particle distribution function $f(\mathbf{w}, t)$;

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + [f, H] = 0 \tag{5.3}$$

where H is the Hamiltonian of the system, and f is the 1-particle distribution function. In a self-consistent system (that is a system which evolves under a

force law which mediates inter-particle interactions) the Hamiltonian takes the form:

$$H[f](\mathbf{w}, t) = \tfrac{1}{2}m\mathbf{v}^2 + \Phi[f](\mathbf{x}, t) \tag{5.4}$$

for self-gravitating systems, the potential Φ is defined via Poisson's equation,

$$\Phi(\mathbf{x}, t) = -GMm \int \mathrm{d}^3\mathbf{w}' \; \frac{1}{|\mathbf{x} - \mathbf{x}'|} f(\mathbf{w}', t) \tag{5.5}$$

Boltzmann's equation can be derived from the BBGKY hierarchy (Binney and Tremaine 2008) which begins with an N-particle distribution function $f^{(N)}(\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_N, t)$, which is reduced to 1-particle distribution functions $f(\mathbf{w}, t)$ under the assumption that no particle is correlated with any other particle.

In general, $f = f(\mathbf{w}, t)$ is a function of phase-space coordinates and time. I have chosen to exclude the additional dependence on time from my notation for the sake of brevity, since it is only rarely involved in the context of time derivatives for which the time-dependence is obvious. Where it appears, the maximum entropy state $f_0(\mathbf{w})$ has no explicit time-dependence.

In this writing, the usual choice of normalisation for the DF f is

$$\int \mathrm{d}^6\mathbf{w} \; f(\mathbf{w}) = 1. \tag{5.6}$$

Astrophysical systems are usually of finite mass, and thus bear that normalisation. However, exceptions exist in the treatment of systems with infinite mass: then, the normalisation is usually done with respect to the density at some scale radii/the spatial density, for homogeneous systems.

5.2.3 Functional Analysis

A functional $G[f](\mathbf{w})$ is in general, an integral over a function of the function f:

$$G[f](\mathbf{w}) = \int \mathrm{d}^6\mathbf{w}_a \; g(f(\mathbf{w}_a), \mathbf{w}). \tag{5.7}$$

The functional integral of a functional $G[f]$ with respect to the function $f(\mathbf{w})$ is formally defined as

$$\int_b^a \mathcal{D}f \; G[f] = \int_b^a ... \int_b^a \Pi_\mathbf{w} \mathrm{d}f(\mathbf{w}) \; G[f] \tag{5.8}$$

where the product goes over all points in phase-space. In physics, this integral is often used to indicate integration over all possible paths or all possible states, where these paths and/or states are defined by f.

A simpler, more intuitive way of arriving at the non-rigorous definition provided above is to consider first a world in which $f(\mathbf{w}) = f_i$ is discretised over some phase-space volume Δ. These phase-space pixels can individually be varied, and all possible f can thus be represented by combinations of all possible f_i.

Then the functional integral is defined as,

$$\int_b^a \mathcal{D}f \; G[f] = \left(\prod_i \int_b^a \mathrm{d}f_i \right) G(f_1, f_2, ...) \tag{5.9}$$

and in the limit $\Delta \to 0$ we reclaim the original result.

The functional derivative obeys the chain rule, product rule, is linear, and is defined by the following relation:

$$\frac{\delta f(\mathbf{w}')}{\delta f(\mathbf{w})} = \delta^6(\mathbf{w} - \mathbf{w}'). \tag{5.10}$$

5.2.4 Generalised Equipartition Theorem

I will now derive a more general variant of the equipartition theorem (Tolman 1938). Consider an abstract Gibbs ensemble in which the probability associated with a state described by $\mathbf{s}$ is;

$$P(\mathbf{s}) = \frac{1}{\mathcal{Z}} \exp\left(-\frac{1}{2}\mathbf{s}^T M \mathbf{s} \right) \tag{5.11}$$

where M is an invertible matrix, summation is implied in the scalar product, and

$$\mathcal{Z} = \int \mathcal{D}\mathbf{s} \; \exp(-\tfrac{1}{2}\mathbf{s}^T M \mathbf{s}) \tag{5.12}$$

is a normalisation factor where the functional integral over $(\mathbf{s})_i = s_i$ goes between $(-\infty, \infty)$. Then ensemble averages with respect to this probability take the form

$$\langle A \rangle = \int \mathcal{D}\mathbf{s} \; P(\mathbf{s})A(\mathbf{s}). \tag{5.13}$$

In this chapter, we are concerned with ensemble averages of functionals A that take the form,

$$\langle s_1 s_1 s_3 ... s_{n(\mathcal{N})} \rangle = \int \mathcal{D}\mathbf{s} \; P(\mathbf{s}) \prod_{i=1}^{\mathcal{N}} s_{n(i)} \tag{5.14}$$

Solving for these $\mathcal{N}$-point correlation functions can be done more simply by considering a related quantity:

$$P(\mathbf{s}, \mathbf{u}) = \exp\left(-\frac{1}{2}\mathbf{s}^T M \mathbf{s} + \mathbf{s}^T \mathbf{u} \right) \tag{5.15}$$

which enables us to write $s_i P(\mathbf{s}, \mathbf{u}) = \partial/\partial u_i P[\mathbf{s}, \mathbf{u}]$. Completing the square, we find:

$$\begin{aligned} P(\mathbf{s}, \mathbf{u}) &= \exp\left(-\frac{1}{2}(\mathbf{s} - M^{-1}\mathbf{u})^T M (\mathbf{s} - M^{-1}\mathbf{u}) \right) \\ &\cdot \exp\left(\frac{1}{2}\mathbf{u}^T M^{-1}\mathbf{u} \right) \end{aligned} \tag{5.16}$$

and then we find that the normalisation factor of this quantity is:

$$\mathcal{Z}(\mathbf{u}) = \mathcal{Z}(0)\exp(\tfrac{1}{2}\mathbf{u}^T M^{-1}\mathbf{u}) \tag{5.17}$$

Utilising equation (5.15) we observe that equation (5.14) can be rewritten as,

$$\left\langle \prod_i s_i \right\rangle = \frac{1}{\mathcal{Z}[0]} \left(\left(\prod_i \frac{\partial}{\partial u_{n(i)}} \right) \mathcal{Z}[\mathbf{u}] \right)\bigg|_{\mathbf{u}=0} \tag{5.18}$$

This reveals the structure of these correlation functions as products of the 2-point correlation functions when $\mathcal{N}$ is even.

$$\left\langle \prod_i s_{n(i)} \right\rangle = \prod_{i,j,i\neq j} M^{-1}_{n(i)n(j)}. \tag{5.19}$$

This result also proves that correlation functions for odd $\mathcal{N}$ go to zero. It also masquerades as the traditional equipartition theorem when we set the state vectors to be positions and velocities in 1D, $\mathbf{s} = (x, v)$ and choose M such that $\frac{1}{2}\mathbf{s}^T M \mathbf{s}$ is the product of the Hamiltonian of a harmonic oscillator and its Lagrange multiplier, $\beta H = \beta(\frac{1}{2}mv^2 + \frac{1}{2}kx^2)$.

These choices produce $\langle x^2 \rangle = 1/k\beta$ and $\langle v^2 \rangle = 1/m\beta$, as expected.

Later in this chapter, the matrix multiplications will be replaced with phase-space integrals, and the partial derivatives with functional derivatives, but the essence of this calculation remains unchanged.

5.3 Probability of probabilities

An N-body system can be summarised as a set of phase-space coordinates, $\mathbf{w}_i = (\mathbf{x}_i, \mathbf{v}_i), i = 1, 2, ..., N$. The equations of motion for such a system are trivial to conceive of–they are $6N$ coupled differential equations, which integrate the particles forward in the (gravitational) potential of their peers. The complete Liouville equation is precisely this,

$$\frac{\mathrm{d} f^{(N)}}{\mathrm{d} t} = \frac{\partial f^{(N)}}{\partial t} + \left[f^{(N)}, H^{(N)} \right] = 0 \tag{5.20}$$

describing the evolution of a distribution of N particles, $f^{(N)}$ under the dynamics prescribed by a corresponding N-particle Hamiltonian $H^{(N)}$.

We want to predict important features in our N-body system; collective effects which are insensitive to exactly the choice of particles comprising a spiral arm, or a Langmuir wave. Such an assertion is equivalent to saying that the particles are sampled in an uncorrelated manner. Then, we have a no-correlations ansatz:

$$f^{(N)}(\{\mathbf{w}_i\}, t) = \prod_{i=1}^{N} f(\mathbf{w}_i, t). \tag{5.21}$$

which states that the N-particle distribution function $f^{(N)}$ is the N-fold product of the one-particle distribution function f at the uncorrelated phase-space coordinates $\{\mathbf{w}_i\}$ and gives us, for large N,

$$\frac{\mathrm{d} f}{\mathrm{d} t} = \frac{\partial f}{\partial t} + [f, H] = 0 \tag{5.22}$$

which is the Collisionless Boltzmann Equation (CBE).

The central assumption we must embrace when we use the CBE is therefore that we must randomly sample our distribution f to obtain coordinates $\{\mathbf{w}_i\}$.

We know that we can obtain almost any set of coordinates $\{\mathbf{w}_i\}$ from the distribution function f insofar as f evaluated at each coordinate is non-zero,

$$f \to \{\mathbf{w}_i\} \text{ if } 0 \notin \{f(\mathbf{w}_i)\} \tag{5.23}$$

but this also means that almost any distribution function f is consistent with sampling the particles $\{\mathbf{w}_i\}$,

$$\{\mathbf{w}_i\} \to f \text{ if } 0 \notin \{f(\mathbf{w}_i)\}. \tag{5.24}$$

We also know that the distribution function f is applied to capture the density of particles, while it is the particles $\{\mathbf{w}_i\}$ that are real! A theory of fluctuations should include finiteness noise. Discarding the preconception that we should fit finite systems with mean field distributions allow us to ask a more fundamental question: What is the optimal distribution function for a sampling of particles, $\{\mathbf{w}_i\}$?

Now, say the optimal distribution function for a sampling of particles $\{\mathbf{w}_i\}$ exists. Pursuing this program, we might first guess that the optimal distribution function is the one which maximises the probability of obtaining $\{\mathbf{w}_i\}$, i.e. $f^{(N)}(\{\mathbf{w}_i\}, t)$, but we soon see that the valid, discrete probability distribution

$$f_d(\mathbf{w}) = \frac{1}{N} \sum_{i=1}^{N} \delta^6(\mathbf{w} - \mathbf{w}_i) \tag{5.25}$$

is a result that captures no new information which singularly maximises the product sum. This result should come at no surprise, as having absolute confidence that the coordinates $\{\mathbf{w}_i\}$ completely defines f relegates us to having to solve the initial problem of $6N$ coupled force equations.

We know that random sampling a distribution to get a sample is an irreversible process, so we shouldn't have confidence (in the Bayesian sense) that any one realisation of $\{\mathbf{w}_i\}$ constrains f in the slightest. I suggest that much in the same way we can sample one choice of f to obtain many $\{\mathbf{w}_i\}$, we should associate with each distribution function, f, a joint probability P_J that a distribution function f was chosen and sample $\{\mathbf{w}_i\}$ was drawn from it, $P_J = P_J[f, \{\mathbf{w}_i\}]$.

$$P_J[f, \{\mathbf{w}_i\}] = P[f] \prod_{i=1}^{N} f(\mathbf{w}_i). \tag{5.26}$$

We know that once a choice of distribution function f is made, we can only randomly sample it to obtain particles $\{\mathbf{w}_i\}$. This implies that P_J can only depend on $\{\mathbf{w}_i\}$ through $\prod_i f(\mathbf{w}_i)$, or that $P = P[f]$ cannot depend on the sample.

5.3.1 Constraining $P[f]$

Constraining $P[f]$ can be done through a maximum entropy argument.

P_J admits a Shannon entropy; which we will denote the joint entropy S_J. Integrating over all distribution functions[2] and all samplings of particles,

$$S_J = \int \mathcal{D}f \int \mathrm{d}^{6N}\mathbf{w} \ (-P_J \ln P_J). \tag{5.27}$$

Substituting equation (5.26) and simplifying leads us to express S_J in terms of $P[f]$ and f,

$$S_J = \int \mathcal{D}f \ (-P[f] \ln P[f] + P[f] S_g[f]) \tag{5.28}$$

where S_g is the Gibbs entropy

$$S_g = -N \int \mathrm{d}^6\mathbf{w} f(\mathbf{w}) \ln f(\mathbf{w}). \tag{5.29}$$

Consequently, S_g may be interpreted as the number of ways in which one might randomly sample N particles from the DF f.

With this foundation set in stone, we can then ask the question: "Given f, how can I choose $P[f]$ so as to maximise my ignorance with regards to the connection between f and $\{\mathbf{w}_i\}$?"

The answer is to calculate $P[f]$ by extremising the combined entropy S_J with respect to $P[f]$ for some known f,

$$\left. \frac{\delta S_J}{\delta P[f]} \right|_f = 0. \tag{5.30}$$

This grants us the unique probability of f being the ignorance-maximising distribution function corresponding to a system which evolves via the CBE:

$$P[f] \propto \exp(S_g[f]) \tag{5.31}$$

---

[2]The anonymous referee pointed out that ensemble averages over the space of all DFs does not respect the normalisation of f. This is true, and this normalisation has to be introduced in an indirect manner, as a constraint on the space of f. This constraint enforces that ensemble averages of f must be normalised.

which upon proper normalisation such that $\int \mathcal{D}f\ P[f] = 1$ via dividing by the partition function, $\mathcal{Z}$, we get:

$$P[f] = \frac{1}{\mathcal{Z}} \exp(S_g[f]). \tag{5.32}$$

The significance of this result is three-fold. First, it confers physical meaning to the Gibbs entropy S_g, concretising its dependence on N oft left neglected. Second, it reveals that the states f with the largest $P[f]$ are also those with the largest $S_g[f]$, appealing to our intuition regarding ignorance: when we know nothing, the best guesses which maximise $P[f]$ also maximises our ignorance regarding how the particles $\{\mathbf{w}_i\}$ are arranged i.e. maximises S_g. Finally, it also shows explicitly that choices of f near the state of maximum Gibbs entropy which we will denote

$$f_0 = f|_{\max(S_g)}, \tag{5.33}$$

have $P[f]$ comparable to $P[f_0]$.

This has the side-effect of ensuring that the ensemble average of f (hereon out, the mean-field), is not f_0 in general,

$$\langle f \rangle = \int \mathcal{D}f\ P[f]f \neq f_0 \tag{5.34}$$

since contributions from near-extremum entropy states also influence the integral, and $P[f]$ is not even about $f = f_0$.

Returning to the original problem of finding the optimal f for some unconstrained sampling of particles $\{\mathbf{w}_i\}$, we can substitute for $P[f]$ from equation (5.26) to obtain P_J which reflects equation (5.30)'s entropy-maximisation condition:

$$P_J[f, \{\mathbf{w}_i\}] = \frac{1}{\mathcal{Z}} \exp(S_g[f]) \prod_i f(\mathbf{w}_i). \tag{5.35}$$

We notice that there is no one optimal choice of f, only a distribution of f which all contribute to the sampling.

It is also evident that S_J, the Shannon entropy of P_J is not wholly maximised. If it were, then we would instead find that P_J is a constant, as that is the unconstrained prior for any probability subject to entropy maximisation.

5.3.2 Shannon's Formulation of Entropy

This then raises the question: What does this partial maximisation of the entropy mean, and why does it matter?

Let us consider an experiement where we only have the sample $\{\mathbf{w}_i\}$ and the knowledge that the sample was randomly sampled. We may think we know nothing about the distribution from which the sample was drawn, but given that we know we are random sampling an unknown distribution function f, we must at the very least be able to discern between different f on the basis of not wanting $\{\mathbf{w}_i\}$ to be an outlier of our guess of f.

Intuitively, this feature of 'outlierliness' in random sampling is captured by comparing the size of the N-particle distribution function (equation (5.21)): $f^{(n)} = \prod_i f(\mathbf{w}_i)$ with the size of the N-particle distribution function of an average sampling of f. If they are of comparable size, then we would say that $\{\mathbf{w}_i\}$ is not an outlier, that we expected $\{\mathbf{w}_i\}$ to be sampled from f.

In this subsection, I will use Shannon's definition of typicality to encode this intuition. More specifically, I will show that $P_J[f, \{\mathbf{w}_i\}]$ as defined in equation (5.35) encodes an assumption: that choices of f for which sampling $\{\mathbf{w}_i\}$ is perfectly typical, are equally to likely to have produced the sample. The joint probability $P_J[f, \{\mathbf{w}_i\}]$ is well-positioned to describe this, since it describes the probability that one samples f, and then samples $\{\mathbf{w}_i\}$ from f.

The N-particle distribution function can be expressed in terms of the sample entropy $S_s[f, \{\mathbf{w}_i\}]$ of a sample

$$S_s = -\sum_i \ln f(\mathbf{w}_i). \tag{5.36}$$

The sample entropy is a measure of how surprising it is that a distribution f produces a sample $\{\mathbf{w}_i\}$. This can be seen from how $-\ln f(\mathbf{w}_i)$ is monotonically increasing as $f(\mathbf{w}_i)$ diminishes. A comparatively large sample entropy

therefore indicates a very lucky sampling (or a wrong guess of f!). Observe:

$$f^{(N)}(\mathbf{w}_i) = \prod_i f(\mathbf{w}_i)$$

$$= \exp\left(\sum_i \ln f(\mathbf{w}_i)\right) \tag{5.37}$$

$$= \exp(-S_s[f, \{\mathbf{w}_i\}]).$$

This result is promising as a measure of outlierliness in random sampling; however it falls short because only a comparatively large sample entropy indicates an outlierly sampling, and we have no measure of what the average sample entropy is.

The answer was found by Shannon (1948), who applied Monte Carlo integration to the sampling entropy. Assuming that $\{\mathbf{w}_i\}$ are sampled from f,

$$\lim_{N\to\infty} \frac{1}{N} S_s[f, \{\mathbf{w}_i\}] = -\int \mathrm{d}^6\mathbf{w} \, f \ln f = \frac{1}{N} S_g. \tag{5.38}$$

where $\{\mathbf{w}_i\}$ are N-length sequences. He then went on to intuit that for finite length samples there must exist some 'normal-looking'-typical samples $\{\mathbf{w}_i\}_T$ that obeyed a similar, but weaker condition owing to the law of large numbers. He showed that for any $\epsilon > 0$ it was possible to find a sample size N such that the entropy of the typical sample is close to the Gibbs entropy in the sense that:

$$\frac{1}{N} S_g[f] - \epsilon < \frac{1}{N} S_s[f, \{\mathbf{w}_i\}_T] < \frac{1}{N} S_g[f] + \epsilon. \tag{5.39}$$

Shannon showed that these typical samples $\{\mathbf{w}_i\}_T$ bore other desirable qualities; for sufficiently large N (or sufficiently small ϵ) the set of typical samples occupies nearly all of probability-space $1-\epsilon$, while comprising only a vanishing fraction of sample-space. This means that almost every random sampling of f of a sufficiently large sample size results in a sample that is typical of f.

It seems reasonable to say that the perfectly typical sample $\{\mathbf{w}_i\}_{\mathrm{PT}}$, which defines the perfectly typical sample entropy, satisfies:

$$S_s[f, \{\mathbf{w}_i\}_{\mathrm{PT}}] = S_g[f]. \tag{5.40}$$

This result defines a sample which is absolutely non-outlierly no matter the choice of ϵ, is a sample which we should expect and a perfectly typical sample. We find that we cannot distinguish between different f for which the sample $\{\mathbf{w}_i\} = \{\mathbf{w}_i\}_{\mathrm{PT}}$ is perfectly typical. Thus we must assign to each $P_J[f, \{\mathbf{w}_i\}_{\mathrm{PT}}] = 1/\mathcal{Z}$ an equal probability of contributing to the sample; reflecting our ignorance. Then we have

$$
\begin{aligned}
P_J[f, \{\mathbf{w}_i\}_{PT}] &= P[f] \exp(-S_s[f, \{\mathbf{w}_i\}_{PT}]) \\
&= P[f] \exp(-S_g[f]) \\
&= 1/\mathcal{Z}
\end{aligned}
\tag{5.41}
$$

that defines $P[f] = \exp(S_g[f])/\mathcal{Z}$, e.g. we reclaim equations (5.32) and (5.35).

5.3.3   No Best Coarse-Grained Distribution

This notion of typicality in random sampling can be used to illustrate many results previously only reachable by intuition. One important result is that there is no way to smooth a sampling of particles $\{\mathbf{w}_i\}$ to obtain a coarse-grained distribution (think a histogram!) function f_{cg} without introducing some kind of constraint/belief of what the grain size should be.

If the volume of each coarse-grained phase-space element is Δ, and the number of particles found within the volume indexed by ℓ is n_ℓ, then the coarse-grained DF f_{cg} is

$$
f_{\mathrm{cg}\ell} = \frac{n_\ell}{N\Delta}.
\tag{5.42}
$$

By definition, all coarse-grained DFs are perfectly typical of their samples (and vice versa):

$$
\begin{aligned}
\prod_i f_{\mathrm{cg}}(\mathbf{w}_i) &= \prod_\ell f_{\mathrm{cg}\ell}^{n_\ell} \\
&= \exp\left(\sum_\ell n_\ell \ln f_{\mathrm{cg}\ell} \right) \\
&= \exp\left(N \sum_\ell \Delta f_{\mathrm{cg}\ell} \ln f_{\mathrm{cg}\ell} \right) \\
&= \exp(-S_g[f_{\mathrm{cg}}]).
\end{aligned}
\tag{5.43}
$$

This means $P_J[f_{\mathrm{cg}}, \{\mathbf{w}_i\}] = 1/\mathcal{Z}$, coarse-grained distribution functions f_{cg} of any Δ have the same joint probability with the sample $\{\mathbf{w}_i\}$.

The space of coarse-grained DFs includes both the discrete distribution function (equation (5.25)) in the limit $\Delta \to 0$ and the constant DF in the opposite limit $\Delta \to \infty$. Observe that each member of this space is assigned equal probabilities of being the correct DF; and are equally likely to be the source of the sample.

Let us think of the $\Delta \to \infty$ limit, which produces a flat DF with an infinitesimal probability density. Sampling this DF produces particles located randomly about phase-space. On the converse side of this problem, let us there is the $\Delta \to 0$ coarse-grained DF, which when sampled only produces stars at fixed points in phase-space.

Inspecting $\{\mathbf{w}_i\}$, how can we possibly know if it was generated by the former DF, or the latter? We cannot: unless we have the ability to gain additional samples of f, so as to gain a better understanding of f, we cannot investigate the veracity of either possibility.

And so we can only assume that all f_{cg}, everything between the flat DF and the Klimontovich DF are equally probable in having generated the sample $\{\mathbf{w}_i\}$, which is the only observable. This result is the cornerstone of this field theory, and establishes the primacy of the Shannon entropy of the DF f (i.e. the Gibbs entropy) as the correct entropy to use in this random-sampling problem.

To put it more simply, this theory begins by acknowledging that samples $\{\mathbf{w}_i\}$ cannot be binned into histograms of the form f_{cg} without choosing a bin-size.

5.3.4   Applying P_J?

P_J is the probability that first, the DF f is chosen, and then a series of samples are sampled. Given that we do not know which f is chosen, but we do know $\{\mathbf{w}_i\}$ it is only sensible to define the sample probability $\mathcal{P}$ by integrating P_J

over f:

$$
\begin{aligned}
\mathcal{P}[\{\mathbf{w}_i\}] &= \int \mathcal{D}f \; P_J[f, \{\mathbf{w}_i\}] \\
&= \int \mathcal{D}f \; P[f] \prod_i f(\mathbf{w}_i) \\
&= \left\langle \prod_i f(\mathbf{w}_i) \right\rangle.
\end{aligned}
\tag{5.44}
$$

$\mathcal{P}[\{\mathbf{w}_i\}]$ is a DF independent measure of the probability of sampling the sample $\{\mathbf{w}_i\}$ that pushes the onus of defining the sample probability onto the constraints imposed on the space of DFs. Invoking a binomial expansion by substituting $f = f_0 + \delta f$, we see we can express the sample probability as an expansion in $\mathcal{N}$-point correlation functions $C_\mathcal{N}$ (that measure the correlations in δf at $\mathcal{N}$ points in phase-space):

$$
\begin{aligned}
\mathcal{P}[\{\mathbf{w}_i\}] &= \prod_i f_0(\mathbf{w}_i) \Bigg(1 + \sum_j \left\langle \frac{\delta f(\mathbf{w}_j)}{f_0(\mathbf{w}_j)} \right\rangle \\
&+ \frac{1}{2!} \sum_{j,k,j \neq k} \left\langle \frac{\delta f(\mathbf{w}_j) \delta f(\mathbf{w}_k)}{f_0(\mathbf{w}_j) f_0(\mathbf{w}_k)} \right\rangle + ... \Bigg) \\
&= \prod_i f_0(\mathbf{w}_i) \Bigg(\sum_{\mathcal{N}=0}^{N} \frac{1}{\mathcal{N}!} \sum_{\{\mathbf{w}_i\}_P \in P(\mathcal{N})} \frac{C_\mathcal{N}[\{\mathbf{w}_i\}_P]}{\prod f_0[\{\mathbf{w}_i\}_P]} \Bigg)
\end{aligned}
\tag{5.45}
$$

where $P(\mathcal{N})$ represents the set of all permutations of selecting $\mathcal{N}$ particles (via their phase-space coordinates) from N particles, non-repeating.

In the construction of P_J, we have assumed that each sampling in $\{\mathbf{w}_i\}$ is sampled independently from each other from f, $\mathcal{P}$ need not reflect random sampling: we manifestly see $\mathcal{N}$-point correlations appearing in the calculation.

Since each sampling of f is independent of the previous one, we must realise that these departures from 'random sampling' arise because f itself has natural internal correlations, which when sampled are imparted onto $\{\mathbf{w}_i\}$.

Later in this chapter, we will see that these $\mathcal{N}$-point correlations are sourced from physical interactions. In the next section, I describe how we can insert constraints (physical or not) into this statistical theory to this effect.

5.4   Entropy Maximisation with constraints

Maximising S_J with respect to multiple constraints

$$\langle G_n[f] \rangle = G_{n0} \tag{5.46}$$

on the space of distribution f can be done via the method of Lagrange multipliers. Let us denote $P_{\{G_n\}}[f]$ as the $P[f]$ which is obtained by maximising

$$S_J - \int \mathcal{D}f \; P[f] \sum_n \beta_n G_n[f] \tag{5.47}$$

where β_n are the Lagrange multipliers, which are defined to maintain the constraints. Then we find:

$$P_{\{G_n\}}[f] = \frac{1}{\mathcal{Z}} \exp\left(S_g[f] - \sum_n \beta_n G_n[f] \right). \tag{5.48}$$

In studies of collisionless systems, we will desire constraints to represent quantities which are globally conserved by the time-evolution of the collisionless Boltzmann equation, i.e. total energy $E[f]$, total angular momentum $\mathbf{L}[f]$, etc. $S_g[f]$ is already conserved as one of the Casimir invariants of the CBE, which are integrals of functions of the DF $g(f)$ which go to zero at infinity rapidly enough to satisfy the last equality.

$$\frac{\mathrm{d}}{\mathrm{d}t} G[f] = \int \mathrm{d}^6\mathbf{w} \; \frac{\partial}{\partial t} g(f) = \int \mathrm{d}^6\mathbf{w} \; [g(f), -H] = 0 \tag{5.49}$$

Notice that when $\{G_n[f]\}$ are conserved by the CBE, that $\mathrm{d}G_n[f]/\mathrm{d}t = 0$ then $\mathrm{d}P_{\{G_n\}}[f]/\mathrm{d}t = 0$, which means this method of assigning a probability to each f is strictly time-independent (even in the nonlinear regime!). This ensures that the ensemble averages obey the CBE if their arguments also obey the CBE, i.e. $\mathrm{d}\langle f \rangle / \mathrm{d}t = 0$ since $\mathrm{d}f/\mathrm{d}t = 0$.

5.4.1   Nonlinear Equilibria

The above statement only delineates that $\langle f \rangle$ is a solution to Boltzmann's equation. A stronger version of this result can be proven with a little more work. Here I will prove that when $\{G_n[f]\}$ are conserved by the CBE, their

mean fields $\langle f \rangle$ are in equilibria; that is that they have no explicit time dependence.

Consider first the system with a mean energy imposed $\langle E \rangle = E_0$ so that $P[f] = P_E[f]$ takes the form

$$P_E[f] = \frac{1}{\mathcal{Z}} \exp(S_g[f] - \beta E[f]). \tag{5.50}$$

$E = E[f]$ is the energy of a system; which is always conserved if the system is isolated and lacks dissipation

$$E[f] = N \int d^6\mathbf{w} \ f(H - \tfrac{1}{2} m \Phi[f]) \tag{5.51}$$

and

$$H(\mathbf{w}) = m(\tfrac{1}{2} \mathbf{v}^2 + \Phi(\mathbf{x})[f] + \Phi_e(\mathbf{x})) = \frac{1}{N} \frac{\delta E[f]}{\delta f(\mathbf{w})} \tag{5.52}$$

is the Hamiltonian of the system. This relationship between the energy and the Hamiltonian (and between the angular momentum of the system and the angular momentum of a particle, as described in the next example) is fully general, since the variation in the energy of a system with respect to the introduction of a particle at a point in phase-space must be the Hamiltonian. I choose a Hamiltonian with self-consistent gravitational potential $\Phi[f]$,

$$\Phi[f] = - \int d^6\mathbf{w}' \ \frac{GM}{|\mathbf{x} - \mathbf{x}'|} f(\mathbf{w}') \tag{5.53}$$

and external potential Φ_e. Note then,

$$
\begin{aligned}
\frac{\partial \langle f \rangle_E}{\partial t} &= \int \mathcal{D}f \ P_E[f] \frac{\partial f}{\partial t} \\
&= \int \mathcal{D}f \ P_E[f]([f, -H]) \\
&= \frac{1}{\beta} \int \mathcal{D}f \ P_E[f]([f, -\beta H]) \\
&= \frac{1}{\beta} \int \mathcal{D}f \ P_E[f] \left(\left[f, \frac{\delta S_g}{\delta f} - \beta \frac{\delta E}{\delta f} \right] \right) \\
&= \frac{1}{\beta} \int \mathcal{D}f \ \left[f, \frac{\delta P_E[f]}{\delta f} \right] \\
&= 0.
\end{aligned}
\tag{5.54}
$$

Between the third and fourth equalities, I have used the fact that $[f, g(f)] = 0$ for any differentiable function g of f, and between the fourth and the fifth equalities, I have used the chain rule. The final equality is found via integration by parts, and the fact that $\delta f(\mathbf{w})/\delta f(\mathbf{w}') = \delta^6(\mathbf{w} - \mathbf{w}')$ is symmetric, but the Poisson brackets are asymmetrical. The boundary terms of the functional integral go to zero because they are constants, and are eliminated by the Poisson brackets.

This result can be extended to show that an ensemble defined by a mean energy and a mean angular momentum is in equilibrium, but only in the corotating frame, where the Hamiltonian in the rotating frame is $\tilde{H} = H + \Omega_p J_z$ owing to the canonical transformation to said rotating frame.

Here, $P[f] = P_{E,L_z}[f]$ takes the form,

$$P_{E,L_z}[f] = \frac{1}{\mathcal{Z}} \exp(S_g[f] - \beta E[f] - \beta_{L_z} L_z[f]) \tag{5.55}$$

If the angular momentum of a system is L_z,

$$L_z[f] = N \int \mathrm{d}^6\mathbf{w} \; f \left(\mathbf{x} - \int \mathrm{d}^6\mathbf{w}' \; f(\mathbf{w}')\mathbf{x}' \right) \times \mathbf{v}. \tag{5.56}$$

then the angular momentum of a particle in the system is:

$$
\begin{aligned}
J_z(\mathbf{w}) &= \left(\mathbf{x} - \int \mathrm{d}^6\mathbf{w}' \; f(\mathbf{w}')\mathbf{x}' \right) \times \mathbf{v} \\
&\quad - \mathbf{x} \times \int \mathrm{d}^6\mathbf{w}' f(\mathbf{w}')\mathbf{v}' \\
&= \frac{1}{N}\frac{\delta L_z[f]}{\delta f(\mathbf{w})}.
\end{aligned}
\tag{5.57}
$$

where the first term of J_z is the angular momentum about the centre of mass, and the second term is the correction term to the angular momentum arising from the centre of mass drift.

I identify the angular speed of the corotating frame as $\Omega_p = \beta_{L_z}/\beta$, and

find that in the corotating frame:

$$
\begin{aligned}
\frac{\partial \langle f \rangle_{E,L_z}}{\partial t} &= \int \mathcal{D}f \; P_{E,L_z} \frac{\partial f}{\partial t} \\
&= \int \mathcal{D}f \; P_{E,L_z}[f]([f, -\tilde{H}]) \\
&= \frac{1}{\beta} \int \mathcal{D}f \; P_{E,L_z}[f] \left([f, -\beta H - \beta_{L_z} J_z] \right) \\
&= \frac{1}{\beta} \int \mathcal{D}f \; P_{E,L_z}[f] \\
&\quad \cdot \left(\left[f, \frac{\delta S_g}{\delta f} - \beta \frac{\delta E}{\delta f} - \beta_{L_z} \frac{\delta L_z}{\delta f} \right] \right) \\
&= \frac{1}{\beta} \int \mathcal{D}f \; \left[f, \frac{\delta P_{E,L_z}[f]}{\delta f} \right] \\
&= 0.
\end{aligned}
\tag{5.58}
$$

It is an elementary exercise to prove that ensembles with a mean drift in the position of the barycenter as well as its velocity, $\mathbf{x}, \mathbf{v}$, in addition to means imposed for E, L_z, are in equilibrium in the corotating, co-drifting frame. A similar extension to Casimir invariants can be inserted between the third and fourth equalities.

Thus we have a non-perturbative result which defines a family of nonlinear equilibria. These equilibria need not be stable to all perturbations; only those induced by the structure of $P[f]$, which are natural to the system. Actually calculating the distributions of these nonlinear equilibria is far more involved, and can be done perturbatively. We investigate how such calculations can be made in the next section.

5.5 Field theory of distributions

The main contribution in Section 5.3 is the representation of correlation functions $C_{\mathcal{N}}$ as contributors to the sampling probability $\mathcal{P}[\{\mathbf{w}_i\}]$ (equation (5.45)). This result describes how correlations raise or lower the probability of observing a constrained N-particle system in a certain configuration.

I will now explain (with a toy model) how we can proceed to calculate these correlation functions, and in doing so open CBE calculations up to the powerful field theory formalism.

5.5.1 Field Theory Basics

A perturbation field theory is very much like any other perturbation theory: it begins by finding a point to Taylor expand around. In this theory, that point is the state of extremised Gibbs entropy, f_0, and this expansion is well-known as the Saddle Point Approximation (of Lagrangian mechanics).

I propose we study a toy model in which I choose where I expand the perturbation theory around. This equivalent to choosing the distribution function of maximum/extremum Gibbs entropy S_g (and in doing so constraining the underlying ensemble of sampled particles). Do note that this choice of constraint is artificial.

$$P_{f_0}[f] = \frac{1}{\mathcal{Z}} \exp\left(\int \mathrm{d}^6\mathbf{w} \ - N f \ln f - \beta_{f_0}(\mathbf{w}) f \right) \tag{5.59}$$

We select β_{f_0} such that,

$$\beta_{f_0}(\mathbf{w}) = \left. \frac{\delta S_g}{\delta f} \right|_{f=f_0} \tag{5.60}$$

to eliminate the first order in the Taylor expansion of equation (5.59), where we expand $f = f_0 + \delta f$. What remains then after we have fixed our choice of f_0, and neglected constants is:

$$P_{f_0}[\delta f] = \frac{1}{\mathcal{Z}} \exp\left(\int \mathrm{d}^6\mathbf{w} \ - \frac{1}{2!} \frac{N \delta f^2}{f_0} + \frac{1}{3!} \frac{N \delta f^3}{f_0^2} + ... \right). \tag{5.61}$$

Notice that we have isolated a Gaussian form out of $P_{f_0}[f]$. The integration that has to occur over all δf however is not between $(-\infty, \infty)$ but rather between $(-f_0, \infty)$, which means that we are not quite able to solve for $\langle A \rangle$ just yet.

Intuition into this issue can be obtained by observing that $P[\delta f = -f_0] = 1/\mathcal{Z} \exp(-N)$, which is vanishingly small compared to the maximum of $P_{f_0}[\delta f] = 1/\mathcal{Z}$ with even $N \approx 50$ or so, indicating that the correction to probability-space where the Gaussian integrand is significant is negligible. $P[f]$ does become large and highly oscillatory at large, negative δf, however this only increases our motivation in using a truncated series in δf to represent $P_{f_0}[\delta f]$, since we know that including more terms in the expansion brings

us further away from the underlying model where the negative f domain is excluded.

So now the ensemble average of some system-wide quantity represented by the functional $A[f]$ is:

$$\langle A \rangle = \int_\infty^\infty \mathcal{D}\delta f \; \frac{1}{\mathcal{Z}} \exp \left(\int \mathrm{d}^6 \mathbf{w} \; -\frac{1}{2!} \frac{N \delta f^2}{f_0} + ... \right) A[\delta f] \tag{5.62}$$

Now consider a theory in which only the first non-trivial term in the argument of the exponent exists; i.e. the quadratic term which then turns the integral into a Gaussian integral, as presented in equation (5.62). This is known as a free theory (Peskin and Schroeder 1995), simply because there are no 'interactions', which are the subsequent cubic/quartic/... terms.

Denoting correlator brackets taken with respect to the free theory with the hyperscript 0, we note that the first non-trivial correlator evaluates to zero, $\langle \delta f \rangle_{f_0}^0 = 0$, due to the even nature of the Gaussian. The second non-trivial correlator may be computed via usage of the equipartition theorem, and takes the form:

$$C_{2,f_0}^0(\mathbf{w}, \mathbf{w}') = \langle \delta f(\mathbf{w}) \delta f(\mathbf{w}') \rangle_{f_0}^0 = \frac{f_0}{N} \delta^6(\mathbf{w} - \mathbf{w}'). \tag{5.63}$$

Higher order correlation functions for the free theory are only products of two-point correlation functions, which are removed via division by the partition function $\mathcal{Z}$. (e.g. Feynman's result that only connected diagrams contribute to the correlations.) Thus the theory is 'free'.

Observe that the two-point correlation function in equation (5.63) is inversely proportional to N, that is that couplings between different points in phase-space are controlled by the size of N—or that N plays the role of a coupling parameter.

The effect of a coupling parameter is more easily examined if the free theory (which usually is defined without any inherent couplings, and is thus 'free'), has no explicit dependence on it. Thus we now change the normalisation of f, defining:

$$\delta \mu = \sqrt{N} \delta f \tag{5.64}$$

and so the newly normalised free theory two-point correlator is

$$\tilde{C}_{f_0}(\mathbf{w}, \mathbf{w}') = \langle \delta\mu(\mathbf{w})\delta\mu(\mathbf{w}') \rangle = f_0 \delta^6(\mathbf{w} - \mathbf{w}'), \tag{5.65}$$

and the equivalent $P_{f_0}[\delta\mu]$ is now:

$$P_{f_0}[\delta\mu] = \frac{1}{\mathcal{Z}} \exp \left(\int d^6\mathbf{w} - \frac{1}{2}\frac{\delta\mu^2}{f_0} + \frac{1}{\sqrt{N}}\frac{1}{3!}\frac{\delta\mu^3}{f_0^2} + ... \right) \tag{5.66}$$

where the new normalisation of the correlator has made explicit the dependence of the interaction terms on the coupling parameters, which are negative integer powers of $\sqrt{N}$.

The existence of couplings beyond the free theory allows us to calculate (small) corrections to it. If the coupling parameter $1/\sqrt{N}$ is sufficiently small, we can apply the expansion,

$$\exp \left(\int d^6\mathbf{w} \frac{1}{\sqrt{N}}\frac{1}{3!}\frac{\delta\mu^3}{f_0^2} \right) \approx 1 + \int d^6\mathbf{w} \frac{1}{\sqrt{N}}\frac{1}{3!}\frac{\delta\mu^3}{f_0^2} + ... \tag{5.67}$$

which produces a series of interaction terms which diminish with increasing order as the expansion parameter $1/\sqrt{N}$.

5.5.2 Feynman Diagrams

Feynman (1949) describes a way to calculate correlators which respect Wick (1950)'s theorem via a diagrammatic approach to calculating combinatorial factors. I will now compute a simple correlation function illustrating this process.

I want to compute the first non-trivial correction to the 1-point correlation function $\langle \delta f(\mathbf{w}) \rangle$ because it is also the first non-trivial correction to the sample probability $\mathcal{P}[\{\mathbf{w}_i\}]$ (equation (5.44)). This can be done by starting with one factor of δf, and then looking for the first interaction term which bumps the number of factors of δf up to an even number, so as to satisfy the generalised equipartition theorem.

In this theory, that is the first order in the cubic expansion we have just

seen. Thus, we may write:

$$\langle \delta f(\mathbf{w}) \rangle \approx \int \mathcal{D}\delta\mu \, \delta\mu(\mathbf{w}) \frac{1}{3!N} \left(\int \mathrm{d}^6 \mathbf{w}_a \frac{\delta\mu_a^3}{f_{0a}^2} \right)$$
$$\cdot \exp\left(-\int \mathrm{d}^6 \mathbf{w} \frac{1}{2} \frac{\delta\mu^2}{f_0} \right) \tag{5.68}$$

This calculation is simplified by the generalised equipartition theorem, which reduces it to calculating:

$$\langle \delta f(\mathbf{w}) \rangle \approx \int \mathrm{d}^6 \mathbf{w}_a \frac{1}{3!N f_{0a}^2} \prod_{\forall \text{Perm.}} \tilde{C}_{f_0}(\mathbf{w}, \mathbf{w}_a) \tilde{C}_{f_0}(\mathbf{w}_a, \mathbf{w}_a) \tag{5.69}$$

where the product goes over all permutations of the two-point correlators. While this combinatorial calculation is simple enough to do by inspection (there are three ways to connect $\mathbf{w}$ to three different $\mathbf{w}_a$s), the number of factors increases factorially, making calculations by inspection untenable.

Feynman describes how to use his diagrams to solve for these combinatorial constants:

$$\langle \delta f(\mathbf{w}) \rangle = \left(\frac{1}{\sqrt{N}} \delta f(\mathbf{w}) \; -\!\!- \times \frac{1}{3!\sqrt{N}} \; -\!\!\!\prec \right)$$
$$= \frac{3}{3!N} \left(\delta f(\mathbf{w}) \; -\!\!\bigcirc \right) \tag{5.70}$$
$$= \frac{3}{3!N} \delta^6(0).$$

In between the first and the second equalities, we have matched up the external vertex $\delta f(\mathbf{w})$ to all three internal vertices (internal being within an integral), producing a combinatorial factor of 3. Then there are no other ways to pair up the remaining two vertices in each case. Finally, the loop integral amounts to $\int \mathrm{d}^6 \mathbf{w}_a \delta^6(0) \delta^6(\mathbf{w}_a - \mathbf{w}) = \delta^6(0)$, completing the calculation.

Now there are two results worth taking note of here. Note $C_1 = \langle \delta f \rangle$ and $C_2 = \langle \delta f \delta f' \rangle$ are of the same leading order in the coupling parameter. This is a 'lucky' feature: C_1's leading order is 0, and its second order is $O(1/N)$. C_2's first order is $O(1/N)$, and so to leading order, they are of the same size in the coupling parameter. Generally however, $C_n \sim O(1/N^{n-1})$.

5.6 Coarse Graining; Discretisation

Perhaps more pertinently, the undefined $\delta^6(0)$ arises as a feature of our calculation. This divergence arises because there are uncountably infinite degrees of freedom in the defining of f while we only have a finite number of particles to constrain f, and can be mastered by discretising the space of distribution functions[3].

Discretising by setting $f(\mathbf{w}_i) = f_i$,

$$f_i = \frac{1}{\Delta} \int_\Delta \mathrm{d}^6 \mathbf{w}\ f(\mathbf{w}) \tag{5.71}$$

as the mean of f within a discrete phase-space element (henceforth, pixel) of volume Δ and $f(\mathbf{w}_i)$ is a constant within that element. This choice of coarse-graining causes $f(\mathbf{w}_i)$ to only change due to fluxes of probability at the boundaries of each pixel, thus decoupling microscopic dynamics happening within a pixel, from the macroscopic observables taken by studying a set of pixels. The correlation functions then map accordingly;

$$\langle \delta f(\mathbf{w}) \delta f(\mathbf{w}') \rangle = \frac{f_0}{N} \delta^6(\mathbf{w} - \mathbf{w}') \to \langle \delta f_i \delta f_j \rangle = \frac{f_{0i}}{N\Delta} \delta_{ij}. \tag{5.72}$$

From hereon out I have chosen to retain the continuum formalism, since Δ is the smallest scale we are interested in.

The regularised first order correction to the expectation of $f(\mathbf{w})$ can be expressed as,

$$\langle \delta f(\mathbf{w}) \rangle \approx \frac{\delta^6(0)}{2N} \to \frac{1}{2N\Delta} \tag{5.73}$$

and the general expansion to order n for the corrections to $\langle \delta f \rangle_f$ is,

$$\langle f(\mathbf{w}) \rangle = f_0(\mathbf{w}) \sum_{i=0}^{n} c_i \left(\frac{1}{N\Delta f_0} \right)^i \tag{5.74}$$

with combinatorial coefficients c_i. Discretisation reveals that the expansion parameter is the inverse of $N_f = N\Delta f_0$, which is the number of particles

---

[3]This divergence can also be mastered by ensuring that ensemble-averaged distribution functions are normalised. The constraint is $\int \mathcal{D}f\ \mu^N = 1$ where $\mu = \int \mathrm{d}^6\mathbf{w}\ f$

expected to lie within a discrete phase-space element from an N-fold sampling of the mode-field.

If N_f is large, then the expansion (for some proper choice of n) approaches asymptotically the real value of $\langle \delta f(\mathbf{w}) \rangle_f$; the statistics we might extract from a gridded 'average N-particle cloud' becomes more refined the more particles lie within each grid, becoming ill-defined when the expected number of particles is less than 1. This arises because the contribution from a distribution function f and sampling of particles $\{\mathbf{w}_i\}$ to the ensemble average (equation (5.32)), $\mathcal{P}[f](\{\mathbf{w}_i\})$ only constrains $P[f]$ if $f(\mathbf{w}_i)$ is non-zero. This causes a massive degeneracy as grids with no particles do not contribute to the ensemble average and do not constrain $P[f]$ whatsoever.

While the Boltzmann equation itself is a continuous probability flow down to and even beyond the level of individual particles, taking the ensemble average introduces divergences which must be regularised via coarse-graining. This necessitates a 'smoothing'-like procedure in which the distribution function is gridded over distances larger than the inter-particle separation. I therefore introduce a statistically motivated method of coarse-graining which is independent of any underlying physics—it only depends on being able to constrain ensemble averages of f.

Under this model, the distribution functions strictly evolve under the CBE; however if and when we decide to compute an observable, we must coarse-grain to obtain ensemble averages. This is one solution to the long-standing issue that no coarse-grained equivalent to the CBE exists.

There is no need for a coarse-grained CBE, only a coarse-grained ensemble average: we cannot fully know what the true f of a system is, given only a finite sample of particles: coarse-graining does not arise in the computation of the dynamics of f, but rather in the computation of the probability density of an observed sample (i.e. an ensemble average).

Further discussion of this coarse-graining regularisation scheme and its impact on the Boltzmann theory will be reserved for the discussion section.

5.7 Mean Energy Theory

Now of chief interest is the imposition of a mean energy constraint, which I denote with the E subscript:

$$\langle E \rangle_{f_0, E} = E_0 \tag{5.75}$$

Setting a mean energy leads us to write,

$$
P_{f_0, E}[f] = \frac{1}{\mathcal{Z}} \exp \Bigg(S_g - \beta E[f] \\
- N \int \mathrm{d}^6 \mathbf{w} \, f(-\ln f_0 + 1 + \beta H_0) \Bigg) \tag{5.76}
$$

where we have redefined $\beta_f(\mathbf{w})$ so as to maintain our choice of the mean field f_0, and $\beta > 0$ means $P[f]$ favours bound states over unbound states.

Expansion in the parameter $\delta\mu$ about $f = f_0 + \frac{1}{\sqrt{N}}\delta\mu$ from equation (5.64) gives us;

$$
P_{f_0, E}[\delta\mu] = \frac{1}{\mathcal{Z}} \exp \Bigg(-\frac{1}{2} \int \mathrm{d}^6\mathbf{w}\,\mathrm{d}^6\mathbf{w}' \, \delta\mu\,\delta\mu' \\
\times \left[\frac{1}{f_0}\delta^6(\mathbf{w} - \mathbf{w}') - \frac{GMm\beta}{|\mathbf{x} - \mathbf{x}'|} \right] \\
+ \int \mathrm{d}^6\mathbf{w} \frac{1}{\sqrt{N}} \frac{1}{3!} \frac{\delta\mu^3}{f_0^2} + ... \Bigg) \tag{5.77}
$$

and the corresponding normalised two-point correlation function $\tilde{C}_{2, f_0, E} = \langle \delta\mu(\mathbf{w})\delta\mu(\mathbf{w}')\rangle_{f_0, E}$ is

$$
\tilde{C}_{f_0, E}(\mathbf{w}, \mathbf{w}') = f_0 \delta^6(\mathbf{w} - \mathbf{w}') \\
+ f_0(\mathbf{w}) f_0(\mathbf{w}') X(\mathbf{x}, \mathbf{x}') \tag{5.78}
$$

where the spatial correlation function $X(\mathbf{x}, \mathbf{x}')$ satisfies the differential equation,

$$
\frac{1}{4\pi A}\nabla^2 X + \int f_0(\mathbf{w})\mathrm{d}^3\mathbf{v} \times X = -\delta^3(\mathbf{x} - \mathbf{x}') \tag{5.79}
$$

and $A = \beta GMm$. One can check via substitution that $\langle \delta\mu(\mathbf{w})\delta\mu(\mathbf{w}')\rangle_{f, E}$ is indeed the inverse of the expression in the square brackets in equation (5.77).

Then the first order correction to $\langle f \rangle_{f_0, E}$ is

$$\langle \delta f(\mathbf{w}) \rangle_{f_0, E} \approx \frac{1}{2} \int \mathrm{d}^6 \mathbf{w}_a \, \frac{1}{N f_0^2}$$
$$\times \tilde{C}_{f,E}(\mathbf{w}_a, \mathbf{w}_a) \tilde{C}_{f,E}(\mathbf{w}_a, \mathbf{w}) \tag{5.80}$$

Where $\tilde{C}_{f,E}(\mathbf{w}, \mathbf{w})$

$$\tilde{C}_{f,E}(\mathbf{w}, \mathbf{w}) = \frac{f_0}{\Delta} \tag{5.81}$$

is finite due to our coarse-graining in phase-space, and the gravitational correlation is set to zero. In the idealised picture of the Boltzmann equation, phase-space is incompressible; what this means is that a high density phase-space element cannot grow in density by gravitationally depleting neighbouring elements of particles; it can only attract other elements of high density together, displacing low density elements in the process. This inability to enrich oneself is what justifies setting the spatial self-correlation to zero, even though the spatial correlations are the strongest for neighbouring position-space elements.

$$\langle \delta f(\mathbf{w}) \rangle_{f_0, E}$$
$$\approx \frac{1}{2} \int \mathrm{d}^6 \mathbf{w}_a \, \frac{1}{N f_0^2} \left(\frac{f_0(\mathbf{w}_a)}{\Delta} \right)$$
$$\times \left(f_0 \delta^6(\mathbf{w} - \mathbf{w}_a) + f_0(\mathbf{w}) f_0(\mathbf{w}_a) X(\mathbf{x}, \mathbf{x}_a) \right)$$
$$= \frac{1}{2} f_0(\mathbf{w}) \left(\frac{1}{N_f(\mathbf{w})} + \int \mathrm{d}^6 \mathbf{w}_a \, \frac{f_0(\mathbf{w}_a)}{N_f(\mathbf{w}_a)} X(\mathbf{x}_a, \mathbf{x}) \right)$$
$$= \frac{1}{2N\Delta} \left(1 + f_0(\mathbf{w}) \int \mathrm{d}^6 \mathbf{w}_a \, X(\mathbf{x}_a, \mathbf{x}) \right) \tag{5.82}$$

where the integral,

$$\mathcal{X}(\mathbf{x}) = \int \mathrm{d}^3 \mathbf{x}_a \, X(\mathbf{x}_a, \mathbf{x}) \tag{5.83}$$

is a solution of the PDE,

$$\left(\frac{1}{4\pi A} \nabla^2 + \int \mathrm{d}^3 \mathbf{v} \, f_0 \right) \mathcal{X}(\mathbf{x}) = -1. \tag{5.84}$$

with boundary conditions (for the homogeneous solution) determined by the boundary conditions of the problem. I will illustrate this with an example in the next section.

5.8 Non-Zero Non-Fluctuations

Having calculated these corrections for the mean energy theory, we can now ask a most crucial question: are these corrections which define the mean-field significant?

If they are, then how we linearise and study the CBE must change, because the ensemble average is modified by terms which are themselves a function of phase-space.

Let us consider what might be the simplest non-trivial system in which these corrections might matter: the self-gravitating Maxwellian. The Maxwellian is well-known as a stable equilibrium: Indeed, we see that there is no need to impose a constraint to 'choose' f_0: Isothermal distributions arise naturally when we expand (e.g. equation (5.50))

$$P_E[f] = \exp(S_g - \beta E) \tag{5.85}$$

about its saddle-point, the isothermal distribution defined by,

$$\left(\frac{\delta S_g}{\delta f} - \beta \frac{\delta E}{\delta f} \right) \Bigg|_{f=f_0} = 0 \implies f_0 \propto \exp(-\beta H_0). \tag{5.86}$$

In particular, I will consider a self-gravitating Maxwellian which is self-consistent within a maximum radius r_m from the origin, and is otherwise 'frozen' outside that radius, much in the same way that Zang (1976) 'froze' the Mestel disc. This can be done by considering the potential imposed on the system from outside $r \leq r_m$ as an external potential.

The maximum entropy state is the Maxwellian, which we have normalised to $\int \mathrm{d}^3\mathbf{v} \, f_0 = \rho_0$, which differs from our usual normalisation of $\int \mathrm{d}^6\mathbf{w} \, f_0 = 1$. This maps $A = \beta G M m \to A = \beta G m$.

$$f_0 = \frac{\rho_0}{(2\pi/\beta)^{3/2}} \exp(-\beta \tfrac{1}{2} m \mathbf{v}^2) \tag{5.87}$$

For this system, we find that the spatial correlation function obtained by substituting equation (5.87) into equation (5.79) is:

$$X(|\mathbf{x} - \mathbf{x}'|) = A \frac{\cos(k_J r)}{r}; \; k_J = \sqrt{4\pi A \rho_0} \tag{5.88}$$

where $k_J = 2\pi/\lambda_J$ is the Jeans wavenumber and λ_J is the Jeans length.

Then $\mathcal{X}(\mathbf{0})$ is simply

$$
\begin{aligned}
\mathcal{X}(\mathbf{0}) &= \int d^3\mathbf{x}' \, X(|\mathbf{x}'|) \\
&= 4\pi \int_0^{r_m} dr \, Ar\cos(k_J r) \\
&= \frac{4\pi A}{k_J^2}\left(k_J r_m \sin(k_J r_m) + \cos(k_J r_m) - 1\right) \\
&= \frac{1}{\rho_0}\left(k_J r_m \sin(k_J r_m) + \cos(k_J r_m) - 1\right)
\end{aligned}
\tag{5.89}
$$

This lets us constrain $\mathcal{X}(\mathbf{x})$, which obeys two boundary conditions. The first is that $\mathcal{X} = \mathcal{X}(r)$, due to the rotational symmetry of the problem. The second is that $\mathcal{X}(0)$ is finite and takes on the calculated value. Applying these constraints gives us,

$$
\mathcal{X}(r) = C_1\frac{\exp(ik_J r)}{r} + C_2\frac{\exp(-ik_J r)}{r} - \frac{1}{\rho_0},
\tag{5.90}
$$

$C_1 = -C_2$ and

$$
\mathcal{X}(0) = C_1(2ik_J) - \frac{1}{\rho_0}.
\tag{5.91}
$$

Then,

$$
\mathcal{X}(r) = \left(k_J r_m \sin(k_J r_m) + \cos(k_J r_m)\right)\frac{\sin(k_J r)}{\rho_0 k_J r} - \frac{1}{\rho_0}
\tag{5.92}
$$

and the first order correction to $\langle\delta f\rangle_{f,E}$ is

$$
\langle\delta f\rangle_{f_0,E} = \frac{1}{2N\Delta}\left(1 + V_v f_0 \mathcal{X}(\mathbf{x})\right) + ...
\tag{5.93}
$$

where V_v is the volume of velocity-space, $V_v = \int d^3\mathbf{v}$. Now define

$$
N_E = \frac{N\Delta\rho_0}{V_v}
\tag{5.94}
$$

and we find that,

$$
\langle\delta f\rangle(\mathbf{w}) = \frac{1}{2N_f}f_0 + \frac{1}{2N_E}f_0(\mu(k_J r_m)\mathrm{sinc}(k_J r) - 1) + ...,
\tag{5.95}
$$

where N_E plays the role of an expansion parameter much like N_f does, and $\mu(x) = x\sin(x) + \cos(x)$ controls the size of the correction. Observe that

$$N_E/N_f = \rho_0/V_v f_0 \ll 1, \tag{5.96}$$

implying that the gravitational interaction has amplified the 'small' $1/N_f$ noise. This is the gravitational dressing of Poisson noise.

This calculation induces a central-peak in the once spatially homogeneous Maxwellian via the sinc function with wavelength equal to the Jeans length, and appears to represent a system which has undergone growth via the linear Jeans instability, and then nonlinearly come to saturation.

That $|\mu(k_J r_m)|$ for most values of r_m grows with r_m supports this hypothesis, since more mass would be drawn under the influence of the Jeans instability. The zeroes of $\mu(k_J r_m)$ are nonlinearly spaced for low radii, but approach a linear spacing for large r_m, $k_J r_m = n\pi$ where is n large, or simply that $2r_m = n\lambda_J$. This appears to indicate that suitably sized self-consistent spheres are capable of holding fluctuating standing waves. These fluctuations evade the ensemble average that can only capture static features which are not removed in the process of averaging.

We can infer that the ensemble average is only sensitive to corrections which are non-zero on average (as we might expect), and this can lead to us not seeing features when they are in fact dynamic and not static.

5.9 Discussion

5.9.1 Typicality and Shannon's Entropy

Perhaps the most questionable (and most interesting) part of this theory lies in its conception: in the construction of the typicality-based calculation with which we justified the maximum entropy approach to obtain $P[f] = \exp(S_g[f])$.

More specifically, in Section 5.3 I conjecture that all f for which an observed sample, $\{\mathbf{w}_i\}$ is perfectly typical are equally likely to be the sampled f. This underpins the field theory.

Entropy maximisation with respect to $P[f]$ does not require that the Gibbs entropy S_g is maximised, and thus does not require the particles to achieve equipartition. I instead describe how our lack of understanding with respect to the 'true' f necessitates that we assume the widest plausible group of f contributes equiprobably to the sample $\{\mathbf{w}_i\}$: that is, the group of f for which $\{\mathbf{w}_i\}$ is always not an outlier are equally likely to be the 'true' f.

This is the only sensible choice to make in the absence of knowledge differentiating between different f, and is what puts the Gibbs entropy S_g above the menagerie of generalised entropies (for the random sampling problem we have here!).

The Gibbs entropy (but in the hands of Shannon) finds its place in the N-particle distribution function—what Shannon would call the probability of randomly sampling a sequence—where it arises as a byproduct of simultaneously taking the logarithm and exponentiating a product of individual sampling probabilities, and then using the law of large numbers. (e.g. compose equations (5.37) and (5.39)).

There there are no axioms of what the Shannon entropy is, no additivity relations to satisfy. It is merely a quantity which measures the expected probability of sampling a sequence, given that each member of the sequence is sampled from f. It defines what is typical, and for sufficiently large N we learn that probability space is dominated by typical samples. It seems that the Shannon entropy is prevalent in physics because real systems which are of large N tend to appear typical, and measurements reflect this.

As mentioned above, we do not need to maximise the Gibbs entropy here. Non-perturbatively, it is integrated over as a part of $P[f]$. Perturbatively, while we expand about f_0, we only do so to facilitate the Saddle Point Approximation. However, one should note that the partition function $\mathcal{Z}$ does not converge if $P[f]$ does not go to zero at $f \to \infty$. This is usually an issue because it signifies that there are states f which do exist that occupy infinite probability $P[f]$. Not so in our case, where f is normalised to 1.

5.9.2 $\langle f \rangle \neq f_0$

The standard program of fitting f_0 to a distribution is flawed when it comes to the study of dynamical systems. Such an assumption is made whenever we try to fit a distribution of stars, or a plasma density, with isothermal f_0, Hernquist disks, Plummer spheres, and even exponential disks! The error is made when we make claims regarding what the distribution function is, as opposed to what it might be.

Instead of making statements regarding the DF, which we have no direct handle on, we should instead be making statements regarding the state of an ensemble of DFs, amongst which we have beliefs regarding which to favour; whether that is a belief regarding a mean energy, or a mean angular momentum, or anything else we would like.

Embracing this inherent uncertainty regarding our knowledge of the DF has to be done through the space of distribution functions, f, which necessitates the use of a field theory: a perturbative field theory in which we have the freedom to vary $f(\mathbf{w})$ at each point in phase-space. Owing to the stochastic sourcing of this deviation, we find that the corrections to f_0 on $\langle f \rangle$ are represented by the expansion parameter $1/N_f$, which describes the Poissonian fractional uncertainty of the number of particles sampled within an element of phase-space, and is also the parameter controlling the size of corrections to the CBE field theory. To leading order, however, these corrections have no phase-space dependence and thus are dynamically irrelevant.

When a mean energy is imposed, another such expansion parameter is produced: $1/N_E >> 1/N_f$, which controls the size of the corrections based on correlations introduced by the Hamiltonian of the system. The relative sizes of these expansion parameters show that the dressing of Poisson noise arises because discreteness introduces fluctuations in phase-space, which are compounded because force calculations between particles separated in position-space sum the force contributions from fluctuations at all velocities within each spatial element (see equation (5.96)).

5.9.3 Nonlinearity

The field theory does not rely on linearising the CBE; and though we split $f = f_0 + \delta f$ it must be understood that each δf is a unique and different f from every other δf; that is we are studying ensembles of independently realised distribution functions, not fluctuations on the same distribution function. The difference here is that different fluctuations on the same system must share the same momentum, energy, et cetera, while different distribution functions can obviously differ in these regards.

This freedom is necessary because of the nature of random sampling. We know that it is almost certain that randomly sampling a stationary, spherical and isotropic distribution function like the isochrone distribution will produce a discrete N-body system which is neither stationary, spherical nor isotropic. Thus we must accept that the reverse is true too: a globular cluster with zero velocity drift could be sampled from a distribution function which does drift in velocities. We must allow ourselves to consider all distribution functions f, of which the vast majority are out of equilibrium.

And yet we see in Section 5.4 that despite this, ensemble averaged distribution functions for systems with physically meaningful constraints are in equilibria. This result formalises the belief that mean-fields should be in equilibria, while embracing the fact that almost all distributions are out of equilibrium.

This is also a proof that isothermal distributions and exponential disks (that is, isothermal and iso-angular-momentum), once dressed in stochastic gravitational fluctuations, are nonlinear equilibria of the CBE, but only in the sense of the mean field being static: meaning there are still time-varying fluctuations 'dancing' around these systems. This is exciting because it illustrates that well-structured noise can preserve the mean-field they surround, and not induce secular evolution.

A perturbative calculation of the self-gravitating Maxwellian reveals that the nonlinearly stable mean field which correponds to it bears signs of having undergone Jeans collapse. I hypothesise that gravitational dressing is made

manifestly calculable through this apparatus! Further confirmation must be done by comparison with simulation/data.

The field theory presents a quasi-nonlinear theory; it is quasi-nonlinear in that we can only expand to some finite order in the expansion, but we know that we are close to the true nonlinear solution.

5.9.4 Collisionality and Coarse-Graining

Regularising this field theory involves coarse-graining phase-space; i.e. taking $\delta^6(0) \to 1/\Delta$.

This represents how a finite number of particles cannot possibly constrain the uncountably infinite degrees of freedom captured by the continuous distribution function. In Section 5.3 I described how it is not possible to choose a coarse-graining grain size without an explicit belief regarding what it should be. This belief was then asserted in Section 5.6 in the form of a minimum grain-size based on the finiteness of N.

The imposition of a minimum grain-size was based on our need to truncate the asymptotic series of $\langle f \rangle$, so as to produce finite, well-defined outputs with the perturbative field theory. This result also makes statistical sense: there is little reason to believe a single particle in a finite phase-space element can constrain the value of f at that point. Note, however, that such a cut is unncessary: taking infinitesimal phase-space elements (i.e. $\Delta \to 0$) is equivalent to taking the strong coupling limit of a Quantum Field Theory, for which perturbation field theory does not apply, since the expansions do not converge.

The concept of coarse-graining phase space given a sample $\{\mathbf{w}_i\}$ so as to reclaim f is not new: see Beraldo e Silva et al. (2019) for a salient analysis on how we can use the Shannon-Nyquist (i.e. anti-aliasing) theorem to set bounds on the fineness of distributions which can be uniquely attributed to a sample. Their analysis concurs with ours: in our notation, roughly they conclude that $1/\Delta \leq N$, or $1/N\Delta \leq 1$ is the limit, where Δ is the phase-space volume of the smallest significant feature they capture. Plasma physicists have also used the Debye sphere, a natural physical scale at which the electrostatic interaction

falls off due to electron shielding, e.g. see Ewart et al. (2023) for a more careful treatment in the context of Lynden-Bell statistics. That considerably larger Δ allows for $1/N\Delta << 1$. But here we should ask ourselves: do we choose the theory which admits as many choices of Δ as possible, or do we select a theory with one, physically motivated Δ?

Perhaps less well-motivated is a presumption I made when I introduced the mean energy constraint in Section 5.8, and then specified that there are no gravitational correlations within the same discrete phase-space element. While this is most certainly true in the continuum limit, since single particles occupying infinitesimal packets of phase-space cannot possibly gravitationally enhance themselves due to the incompressibility of phase-space, discrete phase-space elements contain a finite volume of phase-space, and so infinitesimal packets of phase-space can enter and leave the volume. The discrete phase-space element can thus harbour gravitational enhancements.

What I neglect are the local gravitational collisions which occur between particles within a discrete phase-space element. Such local scatterings are well described by Chandrasekhar (1949). Choosing a size for the discrete phase-space element therefore requires deciding, at what point does local scattering become less important than the Poisson noise one suffers if there are only a few stars within the element? It can be seen that the more we suppress Poisson noise through discretising over larger distances, the more we end up neglecting scattering.

We might understand the relationship between the CBE and local scattering as being one in which the CBE handles all long-distance relationships between the finite phase-space elements, while local scattering handles the local relationships between particles within the phase-space elements. The Coulomb logarithm describes how local scattering logarithmically diverges as one considers scatterings between a source particle and particles which lie successively further from the origin, while Fouvry et al. (2021) describes a logarithmic divergence from global interactions between a source fluctuation and fluctuations

of smaller and smaller scales. They are two extremes of a complete theory of collisions.

This is an avenue for future work, which I am quite eager to pursue.

5.9.5 The Sample Probability

Finally, we have the sample probability, $\mathcal{P}$ (equation (5.45)). Connecting the probability required to sample the sequence of particles, $\{\mathbf{w}_i\}$ with the correlation functions that describe how f is correlated with itself between different points in phase-space, proves that correlations within f at the points $\{\mathbf{w}_i\}$ are equal to the correlations between indistinguishable particles $\{\mathbf{w}_i\}$.

This makes sense: if an ensemble of f bears internal correlations, then an ensemble of samples drawn from the ensemble of f should share the same internal correlations. This is a result which cannot be described by the Liouville equation, which acts on individual f, and cannot incorporate information regarding an ensemble of f.

5.10 Conclusions

I have presented a perturbative field theory which allows us to calculate the mean distribution function, $\langle f \rangle$ and higher moments, with respect to the nonlinear Boltzmann equation.

I show that there is a unique method to assigning probabilities to distribution functions, $P[f]$ for the CBE, conditioned on Shannon's typicality condition. This allows us to calculate two-point correlations, discovering how a distribution function is correlated with itself across phase-space, via an ensemble average.

Defining the ensemble average is important because it allows us to employ the field theory formalism. We are able to explore nonlinear equilibria, calculate the saturated states of certain linear instabilities, and (I think) most importantly, express long-term interactions like global gravitational dressing in a sensible manner.

The next papers in this series will describe how we can use this theory to

describe secular evolution, and present results for rotating systems.

Further work should also include an application to the periodic cube: such an application is in progress, and the periodic cube is promising as an instructional testbed for nonlinear effects. It is currently unclear, but what interests me the most is the connection between this theory and the extensive (but with non-additive entropies) statistical mechanics of Tsallis' (see Tsallis (2011) for a review). In parallel, I think it is worthwhile to attempt introducing collision operators into this currently collisionless theory.

Chapter 6

Concluding Remarks

Each chapter has its own conclusion, so I will write a short passage to round off this thesis.

If there is something to be taken away from reading this, it is that there is a novel theory of statistical mechanics that arises from principles that are radically different from Boltzmann-Gibbs statistical mechanics (BGSM).

This theory defines macroscopic features as features that persist amongst different representative models of a microstate, and allow for the computation of a completely different class of macroscopic feature from what BGSM admits.

It produces predictions which align with those of the BGSM for theories in which the inter-particle interaction is short-ranged, but departs from the classical understanding of the principle of maximum entropy when interactions are long-ranged.

The primacy of the Shannon entropy is established in a theory where particles are Poisson sampled from a density f. It is evident that the generalised entropies (i.e. the Casimir invariants of the CBE) are connected to alternate methods of sampling the system that generate the density f but do not Poisson sample f in populating the system.

Corrections to the traditional entropy-maximising distribution functions are computed and interpreted. These corrections stem from the fact that we do not know how information should be divided between the particles in a realisation of a system and the underlying distribution functions that source

them—maximising the entropy of $S = S[f]$ minimises the information contained within the particles, pushing the onus of defining the system onto f: but this creates predictions that do not reflect the discreteness of the particles. Instead, accounting correctly for this discreteness mixes near-maximum entropy distribution functions into the traditional maximum entropy DF.

This theory also embraces the non-linear nature of the CBE. Instead of treating it as an algebraic equation and solving for non-linearity in that way, it treats the CBE as a functional equation, and handles non-linearity via functional integration. In this way we preserve the symmetry of the CBE without compromise.

There is a lot more that will be done. I am in the act of writing an extension to rotating systems, having completed calculations for it. I have yet to rewrite the paper that constitutes the final chapter, however, to reflect my improved understanding of the relationship between the average microstate, f, and the actual microstate, $\{\mathbf{w}_i\}$.

I have also recently begun expanding towards a theory connecting generalised entropies with astrophysical systems in which there is a distribution of stellar masses, in opposition to the theory where all the stars are indistinguishable from each other. The latter is adequately represented by the Shannon entropy, while the former is not. A preliminary result can be found in a reinterpretation of the Bose gas, that can be understood as a gas of stars in which the presence of larger stars is exponentially suppressed (thus the geometric series).

Of course, this is an unrealistic constraint. In reality the mass distribution of stars at a point in phase-space is not prescribed by Bose-Einstein statistics; but rather by the mechanisms of star formation. The point stands, however, that generalised entropies complete the connection between stars and densities of stars.

The treatment of f as both the probabilistic distribution function and the average microstate is problematic: the average microstate may be com-

prised of greater or fewer particles than the actual microstate observed, causing $\int \mathrm{d}^6\mathbf{w}\, f \neq 1$. The solution to this issue is found when we recall that only ensemble-averaged macroscopic quantities should be normalised; and thus introduce a normalisation constraint. This normalisation constraint resolves many of the issues with the currently released version of the paper.

We have presented a theory that generalises Boltzmann-Gibbs statistical mechanics to allow for long-ranged interactions. Interactions are captured in correlation functions that do not explicitly describe correlations in the sampling of particles (that would violate random sampling), instead correlating the distribution functions that underlie the particles. All this implies is that the true distribution function underlying astrophysical systems are naturally 'lumpy': random sampling a 'lumpy' astrophysical system will produce systems that appear correlated.

While the theory stands alone as an information theoretic construct, it presents an interesting question. Can the correlations that have developed within our galaxy be determined by setting a few low-order constraints (such as that of a mean energy constraint, that only involves functionals that are quadratic in f), or does fitting these correlations require higher-order constraints that specify correlation functions of large orders?

The former would indicate that galaxies are well-modelled as relatively simple systems that have equilibrated under their own gravitational responses, while the latter would indicate that the history of each galaxy—the mergers, the influence of nearby satellites, have had a huge role to play in defining their phase-space structure.

By presenting an overarching mechanism for defining correlation functions of any order, we can begin to compare the one-point correlation (i.e. the observed density of stars) to the two-point correlation, and the two-point to the three-point, et cetera. This will allow us to utilise the observational precision achieved by Gaia to its fullest potential in producing correlations similar to what Toomre and Kalnajs (1991) did in their analysis of their simulations.

Bibliography

T. Antoja, A. Helmi, M. Romero-Gómez, D. Katz, C. Babusiaux, R. Drimmel, D. W. Evans, F. Figueras, E. Poggio, C. Reylé, A. C. Robin, G. Seabroke, and C. Soubiran. A dynamically young and perturbed Milky Way disk. *Nat*, 561(7723):360–362, Sept. 2018. doi: 10.1038/s41586-018-0510-7.

V. A. Antonov. Remarks on the Problem of Stability in Stellar Dynamics. *SovA*, 4:859, Apr. 1961.

V. A. Antonov. On the instability of stationary spherical models with merely radial motions. In *Dynamics of Galaxies and Star Clusters*, pages 139–143, Jan. 1973.

V. Arnold. *Mathematical methods of classical mechanics*, volume 60. Springer, 1989.

E. Athanassoula, M. Romero-Gómez, A. Bosma, and J. J. Masdemont. Rings and spirals in barred galaxies - II. Ring and spiral morphology. *MNRAS*, 400(4):1706–1720, Dec. 2009a. doi: 10.1111/j.1365-2966.2009.15583.x.

E. Athanassoula, M. Romero-Gómez, and J. J. Masdemont. Rings and spirals in barred galaxies - I. Building blocks. *MNRAS*, 394(1):67–81, Mar. 2009b. doi: 10.1111/j.1365-2966.2008.14273.x.

E. Athanassoula, M. Romero-Gómez, A. Bosma, and J. J. Masdemont. Rings and spirals in barred galaxies - III. Further comparisons and links to observations. *MNRAS*, 407(3):1433–1448, Sept. 2010. doi: 10.1111/j.1365-2966. 2010.17010.x.

R. Balescu. Irreversible Processes in Ionized Gases. *Physics of Fluids*, 3(1): 52–63, Jan. 1960. doi: 10.1063/1.1706002.

L. Barbieri, P. Di Cintio, G. Giachetti, A. Simon-Petit, and L. Casetti. Symplectic coarse graining approach to the dynamics of spherical self-gravitating systems. *Monthly Notices of the Royal Astronomical Society*, 512(2):3015–3029, Feb. 2022. ISSN 1365-2966. doi: 10.1093/mnras/stac477. URL `http://dx.doi.org/10.1093/mnras/stac477`.

J. Barnes, J. Goodman, and P. Hut. Dynamical Instabilities in Spherical Stellar Systems. *ApJ*, 300:112, Jan. 1986. doi: 10.1086/163786.

Beraldo e Silva, Walter de Siqueira Pedra, and Monica Valluri. The discreteness-driven relaxation of collisionless gravitating systems: Entropy evolution and the nyquist–shannon theorem. *The Astrophysical Journal*, 872 (1):20, feb 2019. doi: 10.3847/1538-4357/aaf8a7. URL `https://doi.org/10.3847%2F1538-4357%2Faaf8a7`.

J. Binney and C. Lacey. The diffusion of stars through phase space. *MNRAS*, 230:597–627, Feb. 1988. doi: 10.1093/mnras/230.4.597.

J. Binney and R. Schönrich. The origin of the Gaia phase-plane spiral. *MNRAS*, 481(2):1501–1506, Dec 2018. doi: 10.1093/mnras/sty2378.

J. Binney and S. Tremaine. *Galactic Dynamics: Second Edition*. Princeton University Press, 2008.

J. Bland-Hawthorn and T. Tepper-Garcia. Galactic seismology: the evolving "phase spiral" after the Sagittarius dwarf impact. *arXiv e-prints*, art. arXiv:2009.02434, Sept. 2020.

J. Burgers. A mathematical model illustrating the theory of turbulence. volume 1 of *Advances in Applied Mechanics*, pages 171–199. Elsevier, 1948. doi: https://doi.org/10.1016/S0065-2156(08)70100-5. URL `https://www.sciencedirect.com/science/article/pii/S0065215608701005`.

K. M. Case. Plasma oscillations. *Annals of Physics*, 7(3):349–364, July 1959. doi: 10.1016/0003-4916(59)90029-6.

S. Chandrasekhar. Brownian Motion, Dynamical Friction, and Stellar Dynamics. *Reviews of Modern Physics*, 21:383–388, July 1949. doi: 10.1103/RevModPhys.21.383.

S. Chapman and T. Cowling. *The Mathematical Theory of Non-uniform Gases: An Account of the Kinetic Theory of Viscosity, Thermal Conduction and Diffusion in Gases*. Cambridge Mathematical Library. Cambridge University Press, 1990. ISBN 9780521408448. URL `https://books.google.co.uk/books?id=Cbp5JP2OTrwC`.

P.-H. Chavanis. Kinetic theory of long-range interacting systems with angle-action variables and collective effects. *Physica A*, 391:3680–3701, July 2012. doi: 10.1016/j.physa.2012.02.019.

R. Chiba and R. Schönrich. Tree-ring structure of galactic bar resonance. *Monthly Notices of the Royal Astronomical Society*, 505:2412–2426, 06 2021. doi: 10.1093/mnras/stab1094.

R. Chiba, J. K. S. Friske, and R. SchÃ¶nrich. Resonance sweeping by a decelerating Galactic bar. *Monthly Notices of the Royal Astronomical Society*, 500(4):4710–4729, 11 2020. ISSN 0035-8711. doi: 10.1093/mnras/staa3585. URL `https://doi.org/10.1093/mnras/staa3585`.

R. Clausius. Ueber die bewegende kraft der wärme und die gesetze, welche sich daraus für die wärmelehre selbst ableiten lassen. *Annalen der Physik*, 155(3):368–397, 1850.

I. Claydon, M. Gieles, A. L. Varri, D. C. Heggie, and A. Zocchi. Spherical models of star clusters with potential escapers. *MNRAS*, 487(1):147–160, July 2019. doi: 10.1093/mnras/stz1109.

P. Debye and E. Hückel. Zur theorie der elektrolyte. i. gefrierpunktserniedrigung und verwandte erscheinungen. *Physikalische Zeitschrift*, 24:185–206, 1923.

W. Dehnen. Phase-space mixing and the merging of cusps. *Monthly Notices of the Royal Astronomical Society*, 360(3):892–900, 07 2005. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2005.09099.x. URL `https://doi.org/10.1111/j.1365-2966.2005.09099.x`.

J. Dieudonné. *Foundations of modern analysis*. Academic Press, 1969.

C. L. Dobbs, C. Theis, J. E. Pringle, and M. R. Bate. Simulations of the grand design galaxy M51: a case study for analysing tidally induced spiral structure. *MNRAS*, 403(2):625–645, Apr. 2010. doi: 10.1111/j.1365-2966.2009.16161.x.

E. D'Onghia. Disk-stability Constraints on the Number of Arms in Spiral Galaxies. *ApJL*, 808(1):L8, July 2015. doi: 10.1088/2041-8205/808/1/L8.

A. S. Eddington. The distribution of stars in globular clusters. *MNRAS*, 76: 572–585, May 1916. doi: 10.1093/mnras/76.7.572.

A. Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 322(8):549–560, 1905. doi: https://doi.org/10.1002/andp.19053220806. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19053220806`.

J. P. Elliott and D. P. G. *Symmetry in physics*. Macmillan, 1989.

R. J. Ewart, M. L. Nastac, and A. A. Schekochihin. Non-thermal particle acceleration and power-law tails via relaxation to universal lynden-bell equilibria, 2023.

R. P. Feynman. Space-time approach to quantum electrodynamics. *Phys. Rev.*, 76:769–789, Sep 1949. doi: 10.1103/PhysRev.76.769. URL `https://link.aps.org/doi/10.1103/PhysRev.76.769`.

J. B. Fouvry, C. Pichon, J. Magorrian, and P. H. Chavanis. Secular diffusion in discrete self-gravitating tepid discs II. Accounting for swing amplification via the matrix method. *A&A*, 584:A129, Dec. 2015. doi: 10.1051/0004-6361/201527052.

J.-B. Fouvry, C. Hamilton, S. Rozier, and C. Pichon. Resonant and non-resonant relaxation of globular clusters. *Monthly Notices of the Royal Astronomical Society*, 508(2):2210–2225, sep 2021. doi: 10.1093/mnras/stab2596. URL `https://doi.org/10.1093%2Fmnras%2Fstab2596`.

J.-B. Fouvry, C. Hamilton, S. Rozier, and C. Pichon. Resonant and non-resonant relaxation of globular clusters. *arXiv e-prints*, art. arXiv:2103.10165, Mar. 2021.

A. M. Fridman and V. L. Polyachenko. *Physics of gravitating systems, vols I and II.* Springer, 1984.

Gaia Collaboration and A. G. A. e. a. Brown. Gaia Data Release 2. Summary of the contents and survey properties. *A&A*, 616:A1, Aug 2018. doi: 10.1051/0004-6361/201833051.

Gaia Collaboration, A. Vallenari, A. G. A. Brown, T. Prusti, J. H. J. de Bruijne, F. Arenou, C. Babusiaux, M. Biermann, O. L. Creevey, C. Ducourant, D. W. Evans, L. Eyer, R. Guerra, A. Hutton, C. Jordi, S. A. Klioner, U. L. Lammers, L. Lindegren, X. Luri, F. Mignard, C. Panem, D. Pourbaix, S. Randich, P. Sartoretti, C. Soubiran, P. Tanga, N. A. Walton, C. A. L. Bailer-Jones, U. Bastian, R. Drimmel, F. Jansen, D. Katz, M. G. Lattanzi, F. van Leeuwen, J. Bakker, C. Cacciari, J. Castañeda, F. De Angeli, C. Fabricius, M. Fouesneau, Y. Frémat, L. Galluccio, A. Guerrier, U. Heiter, E. Masana, R. Messineo, N. Mowlavi, C. Nicolas, K. Nienartowicz,

F. Pailler, P. Panuzzo, F. Riclet, W. Roux, G. M. Seabroke, R. Sordoørcit, F. Thévenin, G. Gracia-Abril, J. Portell, D. Teyssier, M. Altmann, R. Andrae, M. Audard, I. Bellas-Velidis, K. Benson, J. Berthier, R. Blomme, P. W. Burgess, D. Busonero, G. Busso, H. Cánovas, B. Carry, A. Cellino, N. Cheek, G. Clementini, Y. Damerdji, M. Davidson, P. de Teodoro, M. Nuñez Campos, L. Delchambre, A. Dell'Oro, P. Esquej, J. Fernández-Hernández, E. Fraile, D. Garabato, P. Garc´ia-Lario, E. Gosset, R. Haigron, J. L. Halbwachs, N. C. Hambly, D. L. Harrison, J. Hernández, D. Hestroffer, S. T. Hodgkin, B. Holl, K. Janßen, G. Jevardat de Fombelle, S. Jordan, A. Krone-Martins, A. C. Lanzafame, W. Löffler, O. Marchal, P. M. Marrese, A. Moitinho, K. Muinonen, P. Osborne, E. Pancino, T. Pauwels, A. Recio-Blanco, C. Reylé, M. Riello, L. Rimoldini, T. Roegiers, J. Rybizki, L. M. Sarro, C. Siopis, M. Smith, A. Sozzetti, E. Utrilla, M. van Leeuwen, U. Abbas, P. Ábrahám, A. Abreu Aramburu, C. Aerts, J. J. Aguado, M. Ajaj, F. Aldea-Montero, G. Altavilla, M. A. Álvarez, J. Alves, F. Anders, R. I. Anderson, E. Anglada Varela, T. Antoja, D. Baines, S. G. Baker, L. Balaguer-Núñez, E. Balbinot, Z. Balog, C. Barache, D. Barbato, M. Barros, M. A. Barstow, S. Bartolomé, J. L. Bassilana, N. Bauchet, U. Becciani, M. Bellazzini, A. Berihuete, M. Bernet, S. Bertone, L. Bianchi, A. Binnenfeld, S. Blanco-Cuaresma, A. Blazere, T. Boch, A. Bombrun, D. Bossini, S. Bouquillon, A. Bragaglia, L. Bramante, E. Breedt, A. Bressan, N. Brouillet, E. Brugaletta, B. Bucciarelli, A. Burlacu, A. G. Butkevich, R. Buzzi, E. Caffau, R. Cancelliere, T. Cantat-Gaudin, R. Carballo, T. Carlucci, M. I. Carnerero, J. M. Carrasco, L. Casamiquela, M. Castellani, A. Castro-Ginard, L. Chaoul, P. Charlot, L. Chemin, V. Chiaramida, A. Chiavassa, N. Chornay, G. Comoretto, G. Contursi, W. J. Cooper, T. Cornez, S. Cowell, F. Crifo, M. Cropper, M. Crosta, C. Crowley, C. Dafonte, A. Dapergolas, M. David, P. David, P. de Laverny, F. De Luise, R. De March, J. De Ridder, R. de Souza, A. de Torres, E. F. del Peloso, E. del Pozo, M. Delbo, A. Delgado, J. B. Delisle, C. Demouchy, T. E. Dharmawardena, P. Di Matteo,

S. Diakite, C. Diener, E. Distefano, C. Dolding, B. Edvardsson, H. Enke, C. Fabre, M. Fabrizio, S. Faigler, G. Fedorets, P. Fernique, A. Fienga, F. Figueras, Y. Fournier, C. Fouron, F. Fragkoudi, M. Gai, A. Garcia-Gutierrez, M. Garcia-Reinaldos, M. Garc´ia-Torres, A. Garofalo, A. Gavel, P. Gavras, E. Gerlach, R. Geyer, P. Giacobbe, G. Gilmore, S. Girona, G. Giuffrida, R. Gomel, A. Gomez, J. González-Núñez, I. González-Santamar´ia, J. J. González-Vidal, M. Granvik, P. Guillout, J. Guiraud, R. Gutiérrez-Sánchez, L. P. Guy, D. Hatzidimitriou, M. Hauser, M. Haywood, A. Helmer, A. Helmi, M. H. Sarmiento, S. L. Hidalgo, T. Hilger, N. Hładczuk, D. Hobbs, G. Holland, H. E. Huckle, K. Jardine, G. Jasniewicz, A. Jean-Antoine Piccolo, Ó. Jiménez-Arranz, A. Jorissen, J. Juaristi Campillo, F. Julbe, L. Karbevska, P. Kervella, S. Khanna, M. Kontizas, G. Kordopatis, A. J. Korn, Á. Kóspál, Z. Kostrzewa-Rutkowska, K. Kruszyńska, M. Kun, P. Laizeau, S. Lambert, A. F. Lanza, Y. Lasne, J. F. Le Campion, Y. Lebreton, T. Lebzelter, S. Leccia, N. Leclerc, I. Lecoeur-Taibi, S. Liao, E. L. Licata, H. E. P. Lindstrøm, T. A. Lister, E. Livanou, A. Lobel, A. Lorca, C. Loup, P. Madrero Pardo, A. Magdaleno Romeo, S. Managau, R. G. Mann, M. Manteiga, J. M. Marchant, M. Marconi, J. Marcos, M. M. S. Marcos Santos, D. Mar´in Pina, S. Marinoni, F. Marocco, D. J. Marshall, L. M. Polo, J. M. Mart´in-Fleitas, G. Marton, N. Mary, A. Masip, D. Massari, A. Mastrobuono-Battisti, T. Mazeh, P. J. McMillan, S. Messina, D. Michalik, N. R. Millar, A. Mints, D. Molina, R. Molinaro, L. Molnár, G. Monari, M. Monguió, P. Montegriffo, A. Montero, R. Mor, A. Mora, R. Morbidelli, T. Morel, D. Morris, T. Muraveva, C. P. Murphy, I. Musella, Z. Nagy, L. Noval, F. Ocaña, A. Ogden, C. Ordenovic, J. O. Osinde, C. Pagani, I. Pagano, L. Palaversa, P. A. Palicio, L. Pallas-Quintela, A. Panahi, S. Payne-Wardenaar, X. Peñalosa Esteller, A. Penttilä, B. Pichon, A. M. Piersimoni, F. X. Pineau, E. Plachy, G. Plum, E. Poggio, A. Prša, L. Pulone, E. Racero, S. Ragaini, M. Rainer, C. M. Raiteri, N. Rambaux, P. Ramos, M. Ramos-Lerate, P. Re Fiorentin, S. Regibo,

P. J. Richards, C. Rios Diaz, V. Ripepi, A. Riva, H. W. Rix, G. Rixon, N. Robichon, A. C. Robin, C. Robin, M. Roelens, H. R. O. Rogues, L. Rohrbasser, M. Romero-Gómez, N. Rowell, F. Royer, D. Ruz Mieres, K. A. Rybicki, G. Sadowski, A. Sáez Núñez, A. Sagristà Sellés, J. Sahlmann, E. Salguero, N. Samaras, V. Sanchez Gimenez, N. Sanna, R. Santoveña, M. Sarasso, M. Schultheis, E. Sciacca, M. Segol, J. C. Segovia, D. Ségransan, D. Semeux, S. Shahaf, H. I. Siddiqui, A. Siebert, L. Siltala, A. Silvelo, E. Slezak, I. Slezak, R. L. Smart, O. N. Snaith, E. Solano, F. Solitro, D. Souami, J. Souchay, A. Spagna, L. Spina, F. Spoto, I. A. Steele, H. Steidelmüller, C. A. Stephenson, M. Süveges, J. Surdej, L. Szabados, E. Szegedi-Elek, F. Taris, M. B. Taylo, R. Teixeira, L. Tolomei, N. Tonello, F. Torra, J. Torra, G. Torralba Elipe, M. Trabucchi, A. T. Tsounis, C. Turon, A. Ulla, N. Unger, M. V. Vaillant, E. van Dillen, W. van Reeven, O. Vanel, A. Vecchiato, Y. Viala, D. Vicente, S. Voutsinas, M. Weiler, T. Wevers, L. Wyrzykowski, A. Yoldas, P. Yvard, H. Zhao, J. Zorec, S. Zucker, and T. Zwitter. Gaia Data Release 3: Summary of the content and survey properties. *Astronomy and Astrophysics*, art. arXiv:2208.00211, July 2022.

J. W. Gibbs. *Elementary Principles in Statistical Mechanics*. Yale University Press, 1902.

I. H. Gilbert. Collisional Relaxation in Stellar Systems. *ApJ*, 152:1043, June 1968. doi: 10.1086/149616.

I. H. Gilbert. Gravitational Polarization in Spherical Stellar Systems. *ApJ*, 159:239, Jan. 1970. doi: 10.1086/150306.

P. Goldreich and D. Lynden-Bell. II. Spiral arms as sheared gravitational instabilities. *MNRAS*, 130:125, Jan. 1965. doi: 10.1093/mnras/130.2.125.

R. J. J. Grand, D. Kawata, and M. Cropper. Spiral arm pitch angle and galactic shear rate in N-body simulations of disc galaxies. *A&A*, 553:A77, May 2013. doi: 10.1051/0004-6361/201321308.

M. S. Green. Markoff Random Processes and the Statistical Mechanics of Time-Dependent Phenomena. II. Irreversible Processes in Fluids. *The Journal of Chemical Physics*, 22(3):398–413, 03 1954. ISSN 0021-9606. doi: 10.1063/1.1740082. URL https://doi.org/10.1063/1.1740082.

C. Hamilton. A simple, heuristic derivation of the Balescu-Lenard kinetic equation for stellar systems. *MNRAS*, 501(3):3371–3377, Mar. 2021. doi: 10.1093/mnras/staa3807.

C. Hamilton and T. Heinemann. Noise and waves: a unified kinetic theory for stellar systems. *arXiv e-prints*, art. arXiv:2011.14812, Nov. 2020.

C. Hamilton, J.-B. Fouvry, J. Binney, and C. Pichon. Revisiting relaxation in globular clusters. *MNRAS*, 481:2041–2061, Dec. 2018. doi: 10.1093/mnras/sty2295.

D. C. Heggie, P. G. Breen, and A. L. Varri. l = 1: Weinberg's weakly damped mode in an n-body model of a spherical stellar system. *Monthly Notices of the Royal Astronomical Society*, 492:6019–6026, 2020. URL https://api.semanticscholar.org/CorpusID:211010981.

M. Henon. L'amas isochrone: I. *Annales d'Astrophysique*, 22:126, Feb. 1959.

M. Hénon. Sur l'évolution dynamique des amas globulaires. *Annales d'Astrophysique*, 24:369, Feb. 1961.

M. Henon. Numerical Experiments on the Stability of Spherical Stellar Systems. *A&A*, 24:229, Apr. 1973.

L. Hernquist. An analytical model for spherical galaxies and bulges. *ApJ*, 356: 359–364, June 1990. doi: 10.1086/168845.

J. Heyvaerts. A Balescu-Lenard-type kinetic equation for the collisional evolution of stable self-gravitating systems. *MNRAS*, 407:355–372, Sept. 2010. doi: 10.1111/j.1365-2966.2010.16899.x.

E. Holmberg. On the Clustering Tendencies among the Nebulae. II. a Study of Encounters Between Laboratory Models of Stellar Systems by a New Integration Procedure. *ApJ*, 94:385, Nov. 1941. doi: 10.1086/144344.

J. Jeans. *Problems of Cosmogony and Stellar Dynamics.* Adams prize essay. University Press, 1919. URL `https://books.google.co.uk/books?id=coE-AAAAYAAJ`.

J. H. Jeans. The stability of a spherical nebula. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 199:1–53, 1902. ISSN 02643952. URL `http://www.jstor.org/stable/90845`.

W. H. Julian and A. Toomre. Non-Axisymmetric Responses of Differentially Rotating Disks of Stars. *ApJ*, 146:810, Dec. 1966. doi: 10.1086/148957.

A. J. Kalnajs. *The Stability of Highly Flattened Galaxies.* PhD thesis, Harvard University., 1965.

A. J. Kalnajs. Dynamics of Flat Galaxies. I. *ApJ*, 166:275, Jun 1971. doi: 10.1086/150957.

A. J. Kalnajs. Dynamics of Flat Galaxies. II. Biorthonormal Surface Density-Potential Pairs for Finite Disks. *ApJ*, 205:745–750, May 1976. doi: 10.1086/154330.

A. Khinchin. *Mathematical Foundations of Information Theory.* Dover Books on Mathematics. Dover Publications, 1957. ISBN 9780486604343. URL `https://books.google.co.uk/books?id=0uvKF-LT_tMC`.

I. R. King. The structure of star clusters. III. Some simple dynamical models. *AJ*, 71:64, Feb 1966. doi: 10.1086/109857.

R. Kubo. Statistical-mechanical theory of irreversible processes. i. general theory and simple applications to magnetic and conduction prob-

lems. *Journal of the Physical Society of Japan*, 12(6):570–586, 1957. doi: 10.1143/JPSJ.12.570. URL `https://doi.org/10.1143/JPSJ.12.570`.

R. M. Kulsrud and J. W. K. Mark. Collective Instabilities and Waves for Inhomogeneous Stellar Systems. I. The Necessary and Sufficient Energy Principle. *ApJ*, 160:471, May 1970. doi: 10.1086/150448.

L. Landau. 61 - on the vibrations of the electronic plasma. In D. TER HAAR, editor, *Collected Papers of L.D. Landau*, pages 445–460. Pergamon, 1946. ISBN 978-0-08-010586-4. doi: https://doi.org/10.1016/B978-0-08-010586-4. 50066-3. URL `https://www.sciencedirect.com/science/article/pii/B9780080105864500663`.

P. Langevin et al. Sur la théorie du mouvement brownien. *CR Acad. Sci. Paris*, 146(530-533):530, 1908.

C. F. P. Laporte, I. Minchev, K. V. Johnston, and F. A. Gómez. Footprints of the Sagittarius dwarf galaxy in the Gaia data set. *MNRAS*, 485(3):3134–3152, May 2019. doi: 10.1093/mnras/stz583.

J. Y. Lau. Boltzmann equation field theory i: Ensemble averages, 2023. URL `https://arxiv.org/abs/2310.04508`.

J. Y. Lau and J. Binney. Relaxation of spherical stellar systems. *MNRAS*, 490(1):478–490, Nov. 2019. doi: 10.1093/mnras/stz2567.

J. Y. Lau and J. Binney. Stellar fluctuations of a stellar system. *arXiv e-prints*, art. arXiv:2104.07044, Apr. 2021a.

J. Y. Lau and J. Binney. Modes of a stellar system II: non-ergodic systems. *MNRAS*, 507(2):2562–2567, Oct. 2021b. doi: 10.1093/mnras/stab2175.

J. Y. Lau and J. Binney. Erratum: Relaxation of spherical stellar systems. *MNRAS*, 504(2):2841–2841, June 2021c. doi: 10.1093/mnras/stab1111.

J. Y. Lau and J. Binney. Probabilistic distribution functions. *arXiv e-prints*, art. arXiv:xx, Apr. 2021d.

J. Y. Lau and J. Binney. Modes of a stellar system I: Ergodic systems. *MNRAS*, 507(2):2241–2252, 08 2021a. ISSN 0035-8711. doi: 10.1093/mnras/stab2227. URL `https://doi.org/10.1093/mnras/stab2227`.

Y. Y. Lau and G. Bertin. Discrete spiral modes, spiral waves, and the local dispersion relationship. *ApJ*, 226:508–520, Dec. 1978. doi: 10.1086/156635.

G. Laval, C. Mercier, and R. Pellat. An energy principle for stability. *Nucl.Fus.*, 5:156, 1965.

H. M. Lee and J. P. Ostriker. The Evolution and Final Disintegration of Spherical Stellar Systems in a Steady Galactic Tidal Field. *ApJ*, 322:123, Nov. 1987. doi: 10.1086/165709.

A. Lenard. On bogoliubov's kinetic equation for a spatially homogeneous plasma. *Annals of Physics*, 10(3):390–400, 1960. ISSN 0003-4916. doi: https://doi.org/10.1016/0003-4916(60)90003-8. URL `https://www.sciencedirect.com/science/article/pii/0003491660900038`.

A. Lenard and I. B. Bernstein. Plasma oscillations with diffusion in velocity space. *Phys. Rev.*, 112:1456–1459, Dec 1958. doi: 10.1103/PhysRev.112.1456. URL `https://link.aps.org/doi/10.1103/PhysRev.112.1456`.

C. C. Lin and F. H. Shu. On the Spiral Structure of Disk Galaxies. *ApJ*, 140: 646, Aug. 1964. doi: 10.1086/147955.

B. Lindblad. On the nature of the spiral nebulae. *MNRAS*, 87:420, Mar. 1927. doi: 10.1093/mnras/87.5.420.

C. Maclaurin. *A treatise on fluxions*. London : W. Baynes, 1801.

J. Magorrian. Constraining black hole masses from stellar kinematics by summing over all possible distribution functions. *MNRAS*, 373:425–434, Nov. 2006. doi: 10.1111/j.1365-2966.2006.11054.x.

J. Magorrian. Constraining black hole masses from stellar kinematics by summing over all possible distribution functions. *Monthly Notices of the Royal Astronomical Society*, 373(1):425–434, 10 2006. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2006.11054.x. URL https://doi.org/10.1111/j.1365-2966.2006.11054.x.

D. F. Malin and D. Carter. Giant shells around normal elliptical galaxies. *Nat*, 285(5767):643–645, June 1980. doi: 10.1038/285643a0.

F. Masset and M. Tagger. Non-linear coupling of spiral waves in disk galaxies: a numerical study. *A&A*, 322:442–454, June 1997. doi: 10.48550/arXiv. astro-ph/9902125.

S. D. Mathur. Existence of oscillation modes in collisionless gravitating systems. *MNRAS*, 243:529–536, Apr. 1990.

J. C. Maxwell. Iv. on the dynamical theory of gases. *Philosophical Transactions of the Royal Society of London*, 157:49–88, 1867. doi: 10.1098/rstl.1867. 0004. URL https://royalsocietypublishing.org/doi/abs/10.1098/rstl.1867.0004.

D. E. McLaughlin, J. Anderson, G. Meylan, K. Gebhardt, C. Pryor, D. Minniti, and S. Phinney. Hubble Space Telescope Proper Motions and Stellar Dynamics in the Core of the Globular Cluster 47 Tucanae. *ApJS*, 166(1): 249–297, Sept. 2006. doi: 10.1086/505692.

D. Merritt. Distribution functions for spherical galaxies. *MNRAS*, 214:25P–28P, June 1985.

L. Mestel. On the Galactic Law of Rotation. *Monthly Notices of the Royal Astronomical Society*, 126(6):553–575, 11 1963. ISSN 0035-8711. doi: 10.1093/mnras/126.6.553. URL https://doi.org/10.1093/mnras/126.6.553.

R. W. Michie. On the distribution of high energy stars in spherical stellar systems. *MNRAS*, 125:127, Jan. 1963. doi: 10.1093/mnras/125.2.127.

G. Monari, B. Famaey, A. Siebert, C. Wegg, and O. Gerhard. Signatures of the resonances of a large Galactic bar in local velocity space. *A&A*, 626: A41, Jun 2019. doi: 10.1051/0004-6361/201834820.

P. J. Morrison. . *Phys.Lett.A*, 80:383, Sept. 1980.

P. J. Morrison and B. A. Shadwick. . *Acta Physica Polonica A*, 85:759, Sept. 1994.

C. Navier. , Sur les lois des mouvements des fluides, en ayant égard à l'adhésion des molecules. *Annales de Chimie et de Physique, Paris*, 19:244–260, 1821.

R. W. Nelson and S. Tremaine. Linear response, dynamical friction and the fluctuation dissipation theorem in stellar dynamics. *Monthly Notices of the Royal Astronomical Society*, 306(1):1–21, 06 1999. ISSN 0035-8711. doi: 10.1046/j.1365-8711.1999.02101.x. URL https://doi.org/10.1046/j.1365-8711.1999.02101.x.

L. P. Osipkov. Spherical systems of gravitating bodies with an ellipsoidal velocity distribution. *Pis ma Astronomicheskii Zhurnal*, 5:77–80, Feb. 1979.

P. L. Palmer and J. Papaloizou. Instability in spherical stellar systems. *MNRAS*, 224:1043–1053, Feb. 1987. doi: 10.1093/mnras/224.4.1043.

M. E. Peskin and D. V. Schroeder. An introduction to quantum field theory. 1995. URL https://api.semanticscholar.org/CorpusID:9910245.

G. Piotto, A. P. Milone, L. R. Bedin, J. Anderson, I. R. King, A. F. Marino, D. Nardiello, A. Aparicio, B. Barbuy, A. Bellini, T. M. Brown, S. Cassisi, A. M. Cool, A. Cunial, E. Dalessandro, F. D'Antona, F. R. Ferraro, S. Hidalgo, B. Lanzoni, M. Monelli, S. Ortolani, A. Renzini, M. Salaris, A. Sarajedini, R. P. van der Marel, E. Vesperini, and M. Zoccali. The Hubble Space Telescope UV Legacy Survey of Galactic Globular Clusters. I. Overview of the Project and Detection of Multiple Stellar Populations. *AJ*, 149(3):91, Mar. 2015. doi: 10.1088/0004-6256/149/3/91.

H. C. Plummer. On the problem of distribution in globular star clusters. *MNRAS*, 71:460–470, Mar. 1911.

E. V. Polyachenko, I. G. Shukhman, and O. I. Borodina. Damped perturbations in stellar systems: genuine modes and Landau-damped waves. *MNRAS*, 503(1):660–668, May 2021. doi: 10.1093/mnras/stab537.

V. L. Polyachenko and I. G. Shukhman. General Models of Collisionless Spherically Symmetric Stellar Systems - a Stability Analysis. *SovA*, 25:533, Oct. 1981.

J. J. Ramos and R. L. White. Normal-mode-based analysis of electron plasma waves with second-order Hermitian formalism. *Physics of Plasmas*, 25(3): 034501, Mar. 2018. doi: 10.1063/1.5020794.

F. Roman-Oliveira, F. Fraternali, and F. Rizzo. Regular rotation and low turbulence in a diverse sample of $z \sim 4.5$ galaxies observed with ALMA. *MNRAS*, 521(1):1045–1065, May 2023. doi: 10.1093/mnras/stad530.

N. Rostoker. Superposition of Dressed Test Particles. *Physics of Fluids*, 7(4): 479–490, Apr. 1964. doi: 10.1063/1.1711227.

N. Rostoker and M. N. Rosenbluth. Test Particles in a Completely Ionized Plasma. *The Physics of Fluids*, 3(1):1–14, 01 1960. ISSN 0031-9171. doi: 10.1063/1.1705998. URL https://doi.org/10.1063/1.1705998.

V. S. Safronov. On the gravitational instability in flattened systems with axial symmetry and non-uniform rotation. *Annales d'Astrophysique*, 23:979, Feb. 1960.

P. Saha. Unstable modes of a spherical stellar system. *MNRAS*, 248:494–502, Feb. 1991. doi: 10.1093/mnras/248.3.494.

J. A. Sellwood. The lifetimes of spiral patterns in disc galaxies. *MNRAS*, 410 (3):1637–1646, Jan. 2011. doi: 10.1111/j.1365-2966.2010.17545.x.

J. A. Sellwood and R. G. Carlberg. Transient Spirals as Superposed Instabilities. *ApJ*, 785:137, Apr. 2014. doi: 10.1088/0004-637X/785/2/137.

J. A. Sellwood and R. G. Carlberg. Spiral instabilities: mechanism for recurrence. *MNRAS*, 489(1):116–131, Oct 2019. doi: 10.1093/mnras/stz2132.

J. A. Sellwood and R. G. Carlberg. Spiral instabilities: mode saturation and decay. *Monthly Notices of the Royal Astronomical Society*, 517(2):2610–2621, 10 2022. ISSN 0035-8711. doi: 10.1093/mnras/stac2817. URL `https://doi.org/10.1093/mnras/stac2817`.

J. A. Sellwood and F. D. Kahn. Spiral modes driven by narrow features in angular-momentum density. *MNRAS*, 250:278–299, May 1991. doi: 10.1093/mnras/250.2.278.

C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.

G. G. Stokes. *On the Theories of the Internal Friction of Fluids in Motion, and of the Equilibrium and Motion of Elastic Solids*, page 75–129. Cambridge Library Collection - Mathematics. Cambridge University Press, 1845.

M. Tagger, J. F. Sygnet, E. Athanassoula, and R. Pellat. Nonlinear Coupling of Galactic Spiral Modes. *ApJL*, 318:L43, July 1987. doi: 10.1086/184934.

R. Tolman. *The Principles of Statistical Mechanics*. Oxford University Press, 1938.

J. Tonry and D. P. Schneider. A New Technique for Measuring Extragalactic Distances. *AJ*, 96:807, Sept. 1988. doi: 10.1086/114847.

A. Toomre. On the gravitational stability of a disk of stars. *ApJ*, 139:1217–1238, May 1964.

A. Toomre. Group Velocity of Spiral Waves in Galactic Disks. *ApJ*, 158:899, Dec. 1969. doi: 10.1086/150250.

A. Toomre. What amplifies the spirals. In S. M. Fall and D. Lynden-Bell, editors, *Structure and Evolution of Normal Galaxies*, pages 111–136, 1981.

A. Toomre and A. J. Kalnajs. Spiral Chaos in an Orbiting Patch. In B. Sundelius, editor, *Dynamics of Disc Galaxies*, page 341, 1991.

A. Toomre and J. Toomre. Galactic Bridges and Tails. *ApJ*, 178:623–666, Dec. 1972.

S. Tremaine, M. Henon, and D. Lynden-Bell. H-functions and mixing in violent relaxation. *MNRAS*, 219:285–297, Mar. 1986. doi: 10.1093/mnras/219.2. 285.

C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, 52(1-2):479–487, July 1988. doi: 10.1007/BF01016429.

C. Tsallis. The nonadditive entropy sq and its applications in physics and elsewhere: Some remarks. *Entropy*, 13(10):1765–1804, 2011. ISSN 1099-4300. doi: 10.3390/e13101765. URL `https://www.mdpi.com/1099-4300/13/10/1765`.

J. van de Sande, N. Scott, J. Bland-Hawthorn, S. Brough, J. J. Bryant, M. Colless, L. Cortese, S. M. Croom, F. d'Eugenio, C. Foster, M. Goodwin, I. S. Konstantopoulos, J. S. Lawrence, R. M. McDermid, A. M. Medling, M. S. Owers, S. N. Richards, and R. Sharp. A relation between the characteristic stellar ages of galaxies and their intrinsic shapes. *Nature Astronomy*, 2: 483–488, Apr. 2018. doi: 10.1038/s41550-018-0436-x.

N. G. van Kampen. On the theory of stationary waves in plasmas. *Physica*, 21(6):949–963, Jan. 1955. doi: 10.1016/S0031-8914(55)93068-8.

P. O. Vandervoort. On stationary oscillations of galaxies. *MNRAS*, 339(2): 537–555, Feb. 2003. doi: 10.1046/j.1365-8711.2003.06196.x.

E. Vitral and G. A. Mamon. Does NGC 6397 contain an intermediate-mass black hole or a more diffuse inner subcluster? *A&A*, 646:A63, Feb. 2021. doi: 10.1051/0004-6361/202039650.

M. von Smoluchowski. Zur kinetischen theorie der brownschen molekularbewegung und der suspensionen. *Annalen der Physik*, 326(14):756–780, 1906. doi: https://doi.org/10.1002/andp.19063261405. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19063261405`.

M. D. Weinberg. A Search for Instability in Two Families of Spherical Stellar Models. *ApJ*, 368:66, Feb. 1991. doi: 10.1086/169671.

M. D. Weinberg. Nonlocal and Collective Relaxation in Stellar Systems. *ApJ*, 410:543, Jun 1993. doi: 10.1086/172773.

M. D. Weinberg. Weakly Damped Modes in Star Clusters and Galaxies. *ApJ*, 421:481, Feb 1994. doi: 10.1086/173665.

M. D. Weinberg. Fluctuations in finite-N equilibrium stellar systems. *MNRAS*, 297:101–107, June 1998. doi: 10.1046/j.1365-8711.1998.01456.x.

M. D. Weinberg. Noise-driven evolution in stellar systems - I. Theory. *MNRAS*, 328(1):311–320, Nov 2001. doi: 10.1046/j.1365-8711.2001.04883.x.

G. C. Wick. The evaluation of the collision matrix. *Phys. Rev.*, 80:268–272, Oct 1950. doi: 10.1103/PhysRev.80.268. URL `https://link.aps.org/doi/10.1103/PhysRev.80.268`.

T. A. Zang. *PhD Thesis*. PhD thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1976.