S*pringer* **M***onographs in* **M***athematics*

Chjan Lim Joseph Nebus

Vorticity, Statistical Mechanics, and Monte Carlo Simulation

Springer

Chjan Lim
Department of Mathematical Sciences
Rensselaer Polytechnic Institute
Troy, NY 12180-3590
USA
limc@rpi.edu

Joseph Nebus
Department of Mathematics
National University of Singapore
Singapore
nebusj@rpi.edu

Mathematics Subject Classification (2000): 76-01, 76B47, 82-01, 82B26, 82B80

Library of Congress Control Number: 2006930876

ISBN-10: 0-387-35075-6
ISBN-13: 978-0387-35075-2

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

To Siew Leng and Sean – CCL

To Joseph Francis and Dr Mary Casey – JN

Preface

This book is meant for an audience of advanced undergraduates and graduate students taking courses on the statistical mechanics approach to turbulent flows and on stochastic simulations. It is also suitable for the self-study of professionals involved in the research and modelling of large scale stochastic fluid flows with a substantial vortical component.

Several related ideas motivate the approach in this book, namely, the application of equilibrium statistical mechanics to two-dimensional and 2.5-dimensional fluid flows in the spirit of Onsager [337], and Kraichnan [227], is taken to be a valid starting point, and the primary importance of non-linear convection effects combined with the gravitational and rotational properties of large scale stratified flows over the secondary effects of viscosity is assumed. The latter point is corroborated by the many successful studies of fluid viscosity which limit its effects to specific and narrow regions such as boundary layers, and to the initial and transient phases of the experiment such as in the Ekman layer and spin-up [154] [344].

The main point of applying equilibrium statistical methods to the problems in this book is underscored by the values of the Knudsen number $K = \lambda/l$ (where λ is the mean free path of the molecules of the fluid and l is the smallest relevant macroscopic length scale in the flow) in the body of two-dimensional and 2.5-dimensional large scale fluid flows treated here, namely $K < 10^{-6}$. We further elucidate this point by stressing the fact that in this book, the methods of statistical mechanics are applied not to the fluid as an ensemble of molecules but rather to the flow as an ensemble of vorticity parcels. Nonetheless, many of the techniques used in the statistical treatment of molecular thermodynamics, including the spin-lattice models pioneered in the study of magnetism in condensed matter physics, can be adapted for our primary purpose here.

Our approach of applying equilibrium statistical mechanics to vortical flows centers on the extremization of the free energy $F = U - TS$ where U is the internal energy and S is the entropy. Besides the standard application of Planck's theorem to thermal systems at positive temperatures, where

one minimizes the free energy, we are also interested in vortex problems at negative temperatures, where one maximizes the free energy to obtain the thermodynamically stable statistical equilibria. This point is explored in a simple mean field theory for barotropic flows on a rotating sphere that relates for the first time positive and negative critical temperatures of phase transitions to the key variables of planetary spin, relative enstrophy and kinetic energy.

We note that at low enough positive temperatures T, the minimization of F can be profitably approximated by the easier ground state problem. We further note that the ground state problem by virtue of the minimization of augmented energy functionals, is directly related to steady-state flows of the associated Euler equations. Finally, these special steady-states are related back to decaying Navier-Stokes flows by the Principle of Selective Decay or Minimum Enstrophy, which states that the Dirichlet Quotient (defined as enstrophy over energy) in many damped two-dimensional viscous flows tends asymptotically to a minimum, achieved by the special steady-states.

Vortex statistics is special not only because negative temperatures occur at high energies (a curious phenomenon we will explain in detail), but because it is also characterized by the wide range of temperatures over which extremals of the free energy are close to the corresponding extremals of the internal energy. We will explore the physical reasons for these interesting phenomena in several archetypical examples of vortex dynamics. The most important of these problems are the crystalline or polyhedral equilibria of N point vortices on the sphere, the thermal equilibria of the Onsager vortex gas on the unbounded plane with respect to dynamical equilibria of the rotating two-dimensional Euler equations, and the thermal equilibria at negative temperatures of barotropic vortex dynamics on a rotating sphere.

The works, teachings and direct interactions of many scientists have contributed to this book; it is not possible to mention all their names. We acknowledge in gratitude, Denis Blackmore, Alexandre Chorin, Shui Nee Chow, Peter Constantin, Weinan E, Marty Golubitsky, Leslie Greengard, Tom Hou, Choy Heng Lai, Peter Lax, Andy Majda, Paul Newton, Don Saari, David Sattinger, Junping Shi, Lu Ting, and Lawrence Sirovich who taught one of us kinetic theory.

We would like to thank graduate students Tim Andersen and Xueru Ding for their contributions to graduate seminars at Rensselaer Polytechnic Institute where some of the ideas for this book were tested, and Syed M. Assad for past and continuing collaborations of which some have found their way into this book.

We would like to thank the Department of Mathematical Sciences at Rensselaer Polytechnic Institute and the Departments of Computational Science and of Mathematics at the National University of Singapore for their support of our work.

We thank Achi Dosanjh of Springer Science for initiating and completing the editorial process for this book.

We acknowledge the support of Dr. Chris Arney and Dr. Robert Launer, US Army Research Office and Dr. Anil Deane and Dr. Gary Johnson, US Department of Energy in the form of research grants during the research and writing phase.

Rensselaer Polytechnic Institute, *Chjan C Lim*
National University of Singapore, *Joseph Nebus*
June 2006

Contents

1 Introduction 1
1.1 Statistical Mechanics and Vortex Problems 1
1.2 Euler's Equation for Inviscid Fluid Flow 4

2 Probability 9
2.1 Introduction 9
2.2 Basic Notions 9
2.3 Random Variables and Distribution Functions 10
2.4 Expectation Values and Averages 12
2.5 Variance 15
2.6 Multiple Variables and Independence 17
2.7 Limiting Theorems 19
2.8 Bayesian Probability 25
2.9 Remarks 26

3 Statistical Mechanics 29
3.1 Introduction 29
3.2 Ensembles 29
3.3 Partition Functions 30
3.4 Constraints and Chemical Potentials 32
3.5 Partition Functions by Lagrange Multipliers 34
3.6 Microstate and Macrostates 35
3.7 Expectation Values 38
3.8 Thermodynamics from Z 39
3.9 Fluctuations 42
3.10 Applications 44

4 The Monte Carlo Approach 51
4.1 Introduction 51
4.2 Microstates and Macrostates 51
4.3 Markov Chains 53

4.4 Detailed Balance 55
4.5 The Metropolis Rule 55
4.6 Multiple Canonical Constraints 57
4.7 Ensemble Averages 58
4.8 Random Number Generation 62
4.8.1 Linear Congruence Generators 62
4.8.2 Multiple Recursive and Fibonacci Generators 63
4.8.3 Add-with-Carry and Subtract-with-Carry Generators . . 63
4.8.4 Inverse Congruential Generators 64
4.8.5 Combined Generators 65

5 Spectral Methods 67
5.1 Introduction 67
5.2 Inner Products 67
5.3 Basis Functions 70
5.4 Minimizers 72
5.5 Fourier transforms 73
5.6 Spherical Harmonics 77

6 Discrete Models in Fluids 79
6.1 Introduction 79
6.2 Euler's Equations for Fluid Flows 80
6.3 N-body Hamiltonians 83
6.4 Symplectic Variables 86
6.5 Coordinates and Stereographic Projection 89
6.6 Dynamics on the Plane 91
6.7 Dynamics on the Sphere 103
6.8 Remarks 113

7 Spin-Lattice Models 115
7.1 Introduction 115
7.2 Spin-lattice Models 116
7.3 The Lattice Model 117
7.3.1 Point Strength 118
7.3.2 Normalized Strength 121
7.4 Negative Temperatures 124
7.5 Phase Transitions 125
7.6 Energy-Enstrophy-Circulation Model 127
7.7 Solution of the Spherical Ising Model for $\Gamma = 0$ 128

8 Monte Carlo Simulations 133
8.1 Introduction 133
8.2 Correlation Functions 137
8.3 The Mean Nearest Neighbor Parity 140
8.4 Distances 157

8.5 Remarks . . . 159

9 Polyhedra and Ground States . . . 161
9.1 Introduction . . . 161
9.2 Face-Splitting Operations . . . 162
9.2.1 Centroid Splitting . . . 163
9.2.2 Geodesic Splitting . . . 164
9.2.3 Noncommutivity . . . 165
9.3 Polyhedral Families . . . 168
9.4 Polyhedral Families Versus Vortex Gas . . . 169
9.4.1 Pairwise Interaction Energies . . . 170
9.5 Energy of Split Faces . . . 173
9.5.1 Tetrahedron Splittings . . . 174
9.5.2 130 Vortices . . . 176
9.5.3 Octahedron Splittings . . . 178
9.5.4 258 Vortices . . . 180
9.5.5 Icosahedron Splittings . . . 182
9.5.6 642 Vortices . . . 184

10 Mesh Generation . . . 187
10.1 Introduction . . . 187
10.2 The Vortex Gas on the Sphere . . . 188
10.3 Radial Distribution Function . . . 190
10.4 Vortex Gas Results . . . 192
10.5 Rigid Bodies . . . 206
10.6 Spherical Codes . . . 209

11 Statistical Mechanics for a Vortex Gas . . . 213
11.1 Introduction . . . 213
11.2 The Vortex Gas on the Plane . . . 214
11.2.1 Trapped Slender Vortex Filaments . . . 219
11.3 The Discretized Model . . . 222
11.4 Extremizing E . . . 224
11.5 Variational Problem on the Radius . . . 226

12 Two-Layer Quasi-Geostrophic Models . . . 233
12.1 Introduction . . . 233
12.2 Two-Layer Quasi-Geostrophic Models . . . 233
12.3 Governing Equations . . . 234
12.4 Numerical Models . . . 240
12.5 Numerical Vortex Statistics . . . 241

13 A Coupled Barotropic Flow - Rotating Solid Sphere System 245
13.1 Introduction . . . 245
13.2 The Coupled Barotropic Vorticity – Rotating Sphere System . 246
13.2.1 Physical Quantities of the Coupled Flow – Rotating Sphere Model . . . 247
13.3 Variational Results . . . 249
13.4 Energy-Enstrophy Theory for Barotropic Flows . . . 250
13.4.1 Non-Rotating Extremals . . . 251
13.4.2 Rotating Sphere . . . 253
13.5 Statistical Mechanics . . . 254
13.5.1 Derivation of Spin-Lattice Hamiltonians . . . 254
13.5.2 Gaussian Energy-Enstrophy Model . . . 255
13.6 Spherical Model . . . 256
13.7 Monte Carlo Simulations of the Spherical Model . . . 257
13.8 The Vector Spherical Model for Barotropic Statistics . . . 257
13.9 Remarks . . . 260

References . . . 263

Index . . . 283

1

Introduction

1.1 Connecting Statistical Mechanics to Vortex Problems

The "unreasonable effectiveness of mathematics in the physical sciences," a phrase coined by Eugene Wigner[1], is often mentioned as an exotic property of mathematics. It is an expression of the wonder that models constructed from purely theoretical considerations and reasoned out can predict real systems with great precision. A modern instance of this wonder is found in the writings of Subrahmanyan Chandrasekhar[2]: "In my entire scientific life, the most shattering experience has been the realization that an exact solution of Einstein's equations of general relativity, discovered by Roy Kerr, provides the absolutely exact representation of untold numbers of massive black holes that populate the Universe."

The relationship between mathematics and reality was remarked on even by the Pythagoreans, but was probably not seen as noteworthy to the point of becoming cliché in the earliest era of mechanics. A prediction of the orbits of planets and comets based on Newtonian mechanics and gravitation, for example, could be excellent; but since the laws describing gravitation came from observations of planetary orbits that should be expected. If observation and theory did not agree the theory would not have been used.

It is probably in the laws of gasses that this effectiveness became distracting. The dynamics of a gas can in theory be developed by a Newtonian system, provided one knows to represent it as particle interactions – atomic theory, not generally accepted until the 19th century and still worth debate until Brownian motion was explained – and one knows the laws by which

[1] Eugene Paul Wigner, 1902 - 1995, introduced "parity" to quantum mechanics and discovered the curious properties of the strong nuclear force. [288]

[2] Subrahmanyan Chandrasekhar, 1910 - 1995, was a master of the mechanisms of stars, and predicted the existence of black holes from 1930. [288]

they interact – quantum theory – and one can handle the staggering number of variables needed. These were formidable challenges.

Attempted instead were models based on simple theoretical constraints and few details. Assuming simply that gas was made of particles and these particles moved independently provides the ideal gas law which had been noted and publicized by Robert Boyle, 1627-1691, and Jacques-Alexandre-César Charles, 1746-1823, and others centuries earlier. For such a simple model matching the general behavior of real gases was a surprise. Adding the assumptions atoms had a minimum size and some interaction allowed Johannes Diderik van der Waals, 1837-1923, to offer in 1873 a correction to the ideal gas law and an even better match to observation [25] [64].

Erwin Schrödinger[3] [389], in his compelling 1944 essay "What is life?", presented a famous argument for the apparent exactness and determinateness of macroscopic laws which are nonetheless based on physical laws of a statistical nature for the detailed components of a system. He argued the macroscopic law of bulk diffusion (which is clearly deterministic) is based on the completely and purely statistical phenomenon of random walks at the microstate level. The random walk of molecules is not only statistical in nature but it is also completely symmetrical: a molecule takes a jump to the right or left in say a one-dimensional model with equal probability. Yet its consequence at the macroscopic level is clearly asymmetric because the law of bulk diffusion has a clear direction: from high concentration to low concentration.

This is our textbook's inspiration. A simple model of fluid flow will be made from theoretical considerations. The model will be studied through several alternative strategies and adjusted to make it more natural.

We are more concerned with statistical equilibria than with dynamical equilibria. A dynamical equilibrium requires the components of a state to be in a spatially rigid and temporally stationary relationship with each other. This is too restrictive for the problems in this book. Statistical equilibria have stationary macroscopic variables which offer vastly more degrees of freedom in the fine details of the state. A rule of thumb for the appearance of seemingly exact macroscopic laws is that the macroscopic system must have large enough number N of microscopic components in order for the fluctuations of size $1/\sqrt{N}$ in the macroscopic variable to be small.

This book is our attempt to connect two main topics of asymptotic states in vortex flows and equilibrium statistical mechanics. While fully developed turbulence in a damped driven flow is a non-equilibrium phenomena, many powerful arguments (by Kolmogorov [223], [224], Oboukhov [330]) have been presented, asserting that for certain inertial ranges in the power spectrum of driven viscous flows, the methods of equilibrium statistical mechanics can be adopted. We will avoid such arguments and treat the phenomena of isolated

[3] Erwin Rudolf Josef Alexander Schrödinger, 1887 - 1961, got the inspiration for the wave form of quantum mechanics from a student's suggestion at a seminar Schrödinger gave on the electron theory of Louis de Broglie, 1892 - 1987. [288]

inviscid fluid turbulence within the context of equilibrium statistical mechanics.

The concept of negative temperature was introduced into vortex dynamics by Lars Onsager. Vortical systems in two dimensions and in 2.5 dimensions (which we will describe) support negative temperatures at high kinetic energies where the thermal equilibria are characterized by highly organized large-scale coherent structures. Thus, besides the standard application of Planck's theorem to thermal systems at positive temperatures, where one minimizes the free energy, we are also interested in vortex problems at negative temperatures, where one maximizes the free energy to obtain stable statistical equilibria.

In addition to the first common theme of Monte Carlo simulations of organized and of turbulent fluid flows in this book, a second theme is the relationship between dynamics and equilibrium statistical mechanics: the **extremals** (maxima and minima) of the energy determine the equilibria, but the extremals of the free energy give us the most probable states of the equilibrium statistics.

Vortex statistics has many noteworthy examples where the range of temperatures is quite large, over which extremals of the free energy are close to the corresponding extremals of the internal energy. We will explore the physical reasons for these interesting phenomena in several archetypical examples of vortex dynamics. The most important of these problems are the crystalline or polyhedral equilibria of N point vortices on the sphere, the thermal equilibria of the Onsager vortex gas on the unbounded plane with respect to dynamical equilibria of the rotating two-dimensional Euler equations, and the thermal equilibria at negative temperatures of barotropic vortex dynamics on a rotating sphere.

In the first case, N similar point vortices on a sphere, the Monte Carlo simulator running at positive temperatures achieves thermal equilibria which are very close to the polyhedral relative equilibria of the dynamical equations. These polyhedral crystalline states have extremely regular and uniform spatial separations, and thus minimize the interaction energy though without simultaneously maximizing the entropy. This situation provides one of the canonical ways in which minimizers of the free energy are close to the dynamical equilibria for a range of positive temperatures.

The second case concerns the unbounded Onsager point vortex gas, whose thermal equilibria of uniform vorticity distributions in a disk are close to the dynamic equilibria of the rotating two-dimensional Euler equations over a wide range of positive temperatures. The physical reason in this case is the same as the first, that is, the free energy minimizers are given by vortex states which minimize the internal energy.

The physical reason for the third case is that free energy maximizers corresponding to stable thermal equilibria at negative temperatures are achieved by vortex states with very low entropy. Unlike standard thermodynamic applications where entropy is maximized, the solid-body rotation flow states have

the minimum entropy and maximum kinetic energy over allowed flow states with the same relative enstrophy.

A non-extensive continuum limit is allowed for two-dimensional flows in a fixed finite domain. The canonical examples for such flow domains are the fixed bounded regions on the plane, and finite but boundary-less domains such as the surface of the sphere or of the torus. Finite boundary-less domains are computationally convenient because boundary conditions are often complex. And among finite, boundary-less domains the problem of flows on the sphere is more important than flows on more topologically complex surfaces because of their applicability to atmospheric sciences, to which we will return.

Our principal focus is the of inviscid fluids. We justify this choice – which seems to exclude most of the fluids of the real world – on several grounds. The first is that often the viscosity of a fluid is a minor effect, and these slightly viscous fluids can be modelled by inviscid fluids, where we represent the interior of the flow by a fluid without viscosity, and add boundary layers in the regions where viscosity becomes relatively significant. More, even if we want to consider viscous fluids, we can still represent an interesting aspect of them – the non-linear convective aspects of the flows – by treating this portion of the flow as an inviscid fluid.

And furthermore much of what we can study in the thermal equilibrium of inviscid two-dimensional vortex dynamics (such as the minimizers of free energy functions) can be extended naturally to the ground states of augmented energy functionals, or to the steady states of two-dimensional Euler equations. These in turn are related, by the Principle of Selective Decay, also termed the Principle of Minimum Enstrophy, to the asymptotic flows of the decaying two-dimensional Navier[4]-Stokes[5] equations. The minimizers of the Dirichlet quotient, the ratio of enstrophy to energy, corresponds to the inviscid steady states we explicitly study [26].

1.2 Euler's Equation for Inviscid Fluid Flow

To write Leonhard Euler's[6] equation for fluid flow, we begin with the fluid velocity. Letting $\mathbf{u}$ stand for the velocity and ρ the density of the fluid, we

[4] Claude Louis Marie Henri Navier, 1785 - 1836, was in his day famous as a builder of bridges. He developed the first theory of suspension bridges, which had previously been empirical affairs. [288]

[5] George Gabriel Stokes, 1819 - 1903, besides the theory of fluid flow and the theorem about the integrals over surfaces, provided the first theory and the name for fluorescence. [288]

[6] It is almost impossible to overstate the contributions of Leonhard Euler, 1707 - 1783. The scope is suggested by the fact after his death it required fifty years to complete the publication of the backlog of his papers. He is credited with the modern uses of the symbols e, i, π, Σ, the finite differences Δt and $\Delta^2 t$, and $f(x)$ as a general symbol for a function. [288]

choose some fluid properties. We want the fluid to be incompressible, inviscid, and to experience no outside forces.

The obviously important properties of the fluid are the density at a time t and a point $\mathbf{r}$ – call that $\rho(t, \mathbf{r})$ – and the velocity, again a function of time and position. Call that $\mathbf{u}(t, \mathbf{r})$. We will build on three properties [88].

First is the conservation of mass. Suppose the fluid is incompressible, which is nearly correct for interesting fluids such as water at ordinary temperatures and pressures. Incompressibility demands a divergence of zero:

$$\boldsymbol{\nabla} \cdot \mathbf{u} = 0 \tag{1.1}$$

A nonzero divergence over some region A corresponds to either a net loss or net gain of mass, so the fluid density is changing and the fluid is either expanding or compressing.

The next property is the conservation of momentum. The momentum inside region A, the total of mass times velocity, will be

$$\mathbf{p} = \int_A \rho \mathbf{u} dV \tag{1.2}$$

(with dV the differential volume within the region A). So the rate of change of the momentum in time will be

$$\frac{\partial}{\partial t}\mathbf{p} = \int_A \frac{\partial}{\partial t}(\rho \mathbf{u}) dV \tag{1.3}$$

Without external pressure, or gravity, or viscosity or intermolecular forces the momentum over the region A cannot change on the interior. Only on the surface can momentum enter or exit A:

$$\frac{\partial}{\partial t}\mathbf{p} = \int_{\partial A} (\rho \mathbf{u})\mathbf{u} \cdot \mathbf{dS} \tag{1.4}$$

with ∂A the surface of A and $\mathbf{dS}$ the differential element of area for that surface. Using Green's theorem[7], the integral is

$$\frac{\partial}{\partial t}\mathbf{p} = -\int_A \rho(\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{v} dV \tag{1.5}$$

If there is a force, which we will generalize by calling the pressure and denoting it as $P(\mathbf{r}, r)$, then momentum may enter or exit the region A, but again only through its surface. Even more particularly only the component of the force which is parallel to the outward unit normal vector $\hat{n}$ can affect the fluid, in or out. So the change in the momentum of the fluid caused by the pressure term P, is

[7] George Green, 1793 - 1841, besides his theorem connecting surface integrals to volume integrals, is credited with introducing the term "potential function" in the way we use it today. [288]

$$\frac{\partial}{\partial t}\mathbf{p} = \int_{\partial A} P\hat{n}\mathbf{dS} \tag{1.6}$$

$$= -\int_{A} \boldsymbol{\nabla} P\mathbf{dV} \tag{1.7}$$

using again Green's theorem.

As we may have momentum gained or lost through either the fluid flow or through the pressure we add the two terms:

$$\frac{\partial}{\partial t}\mathbf{p} = -\int_{A} \rho(\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{v}dV - \int_{A} \boldsymbol{\nabla} P\mathbf{dV} \tag{1.8}$$

Between equations 1.3 and 1.8 we have two representations of the derivative of momentum with respect to time. Setting them equal

$$\frac{\partial}{\partial t}\mathbf{p} = \int_{A} \frac{\partial}{\partial t}(\rho\mathbf{u})dV = \int_{A} \rho(\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{u}dV - \int_{A} \boldsymbol{\nabla} P\mathbf{dV} \tag{1.9}$$

for all regions A. For the middle and right half of equation 1.9 to be equal independently of A requires the integrands be equal[8]:

$$\frac{\partial}{\partial t}(\rho\mathbf{u}) = -\rho(\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{u} - \boldsymbol{\nabla} P \tag{1.10}$$

which is Euler's equation for inviscid, incompressible, unforced fluid flow. Assuming incompressibility makes ρ constant in time, so we may divide it out.

Having introduced the pressure, we will proceed now to drop it for nearly the entirety of the book, as we will find abundant interesting material even before adding pressure to the system. In this form and confined to one spatial dimension is often known as Burgers' equation[9], though we will keep a bit more freedom in space:

$$\frac{\partial}{\partial t}\mathbf{u} + \mathbf{u}\cdot\boldsymbol{\nabla} u = 0 \tag{1.11}$$

Up to this point we have considered only two important physical properties. The third we will add in order to convert this equation into a form more suitable for treatment as a particle problem, which we will do in chapter 6. There we will also change our attention from the velocity of the fluid into the vorticity, that is, the curl of the velocity. This combination lets us recast the flow of an inviscid fluid as a statistical mechanics problem.

[8] Strictly speaking, they must be equal "almost everywhere" – the set of points that are exceptions must have measure zero. For example, finitely many exceptions are allowed.

[9] Johannes Martinus Burgers, 1895 - 1981. Burgers is known also for the Burgers dislocation, a method of describing the irregularities in a crystal. He was also expert on the study of polymers, and a sort of viscoelastic material is named for him.

Our roots in statistical mechanics and thermodynamics suggests a question: is there a temperature to a vortex dynamical system? Statistical mechanics defines the temperature of any system to be the derivative of energy with respect to the entropy. In the kinetic theory of gases this equals the physical heat. Although there is no physical heat in this problem, there is an energy and there is an entropy; therefore, it has a temperature.

This extension of temperature is not unique. Most physical models have an energy. The entropy of a model can be given through information-theoretical methods – as long as a system can contain information, it has an entropy. Therefore the idea of temperature can be applied to systems that have no resemblance at all to the gas particles the idea began with.

One fascinating consequence is that vortex systems can have a negative temperature. There are configurations for which adding more energy will decrease the entropy of the system. The derivative is then negative and the temperature is therefore less than zero. More remarkably these negative temperature states are extremely high-energy ones. These negative temperature states will receive considerable attention. (Vortex dynamics is not the only context in which negative temperatures arise. They can develop in systems in which a maximum possible energy exists. One noteworthy example is in describing the states of a laser.)

We will need to simplify our problem to be able to apply statistical mechanics methods to it. We want a large but finite number of particles or lattice sites which obey some interaction law. Our interests will lead us to rewrite the Euler equation from several perspectives. In one we will describe the vorticity of the fluid as a set of discrete "charged" particles which are free to move. In another we will construct a piecewise-continuous approximation to the vorticity based around a fixed set of mesh sites and allow the value of the function on these pieces to vary.

If we are interested in the "vortex gas" problem, placing a set of vortices of fixed strength and allowing them to move, then we could write it as a dynamical systems problem, with a Hamiltonian[10], a representation using the form of classical mechanics. With that we can use tools such as the Monte Carlo Metropolis Rule to explore this space and study the equilibrium statistical mechanics.

Unfortunately the Monte Carlo study of the vortex gas problem does not well handle vortices of positive and negative strengths mixed together. The Metropolis-Hastings rule will tend to make vortices of opposite sign cluster together. Similarly negative temperatures cannot be meaningfully applied; trying simply causes all like-signed vortices to cluster together. But as long as

[10] These functions were introduced by Sir William Rowan Hamilton, 1805 - 1865, and have become a fundamental approach to dynamical systems. Hamilton also discovered quaternions, famously carving the inspired equation $i^2 = j^2 = k^2 = ijk = -1$ into the stones of the Brougham Bridge. [288]

we are interested in a single sign and positive temperatures interesting work may be done.

In the lattice problem (our mesh sites may not be the regularly organized rows and columns of a proper lattice, but it is a fixed set of sites) we approximate the continuous vorticity field by a piecewise-continuous approximation. Changes in the fluid flow are represented by changes in the relative strengths of lattice sites. This approach resembles strongly a finite-elements study. This approach also well handles both positive and negative vorticities, and both positive and negative temperatures are meaningfully studied.

There is also a useful approach not based on points and site vorticities at all. Anyone who has studied enough differential equations has encountered Fourier decompositions of problems – supposing that the solution to a differential equation is the sum of sine and cosines of several periods, and finding the relative amplitudes of the different components. This sort of approach is called the spectral method. The analogy to identifying the components of a material by the intensities at different frequencies of the spectrum of light that has passed through the material is plain.

Through these approaches, we plan to show how analytical and computational mathematics complement one another. Analytic study of fluid flow provides a problem well-suited to numerical study. Numerical experiments will improve the understanding of old and will inspire new analysis. In combination we make both approaches stronger.

To end this introduction, we remember that the astronomical and cosmological examples alluded to above have a different scale of predictability than fluid phenomena such as flow turbulence and the weather. Astronomers have predicted solar and lunar eclipses to the second for centuries. But the weather cannot even now be predicted with anywhere near the same scale of accuracy. The accuracy in astronomical prediction is largely dependent on the exactness of the initial data at an earlier epoch. The inaccuracies in weather prediction persists in spite of greatly improved meteorological methods and instruments for measuring the state of an atmosphere. These are two very different realms of applied and computational mathematics, underscoring the theoretical and technical difficulties of the latter.

2

Probability

2.1 Introduction

In science fiction writer Stanley G Weinbaum's 1935 short story "The Lotus Eaters" the ultimate law of physics is declared to be the law of chance. Given the overwhelming power of statistical mechanics and of quantum mechanics this assessment is hard to dispute. The study of probability combines beautiful reasoning beginning from abstract first principles and describes the observable world with remarkable clarity. So before moving into Monte Carlo methods it is worthwhile reviewing the fundamentals of probability.

2.2 Basic Notions

Let the letter S represent a **sample space**, which is the set of all the possible outcomes of a particular experiment. An **event** U is a subset of the sample space: $U \subseteq S$. We typically think of an event as just one of the possible outcomes, but we will be interested in the probability of several outcomes occurring together, or of one outcome occurring given that another has.

We want to define is the **probability** $P\{E\}$ of the event E occurring. The first thought is to do this by a limit. Let the number $n(E)$ represent the number of times outcome E occurs in the first n repetitions of an experiment. Then

$$P\{E\} = \lim_{n\to\infty} \frac{n(E)}{n} \tag{2.1}$$

which is known as the **limiting frequency** of E.

As often happens, this intuitive definition requires the assumption of several difficult-to-justify axioms. The first troublesome axiom is that the limit described above exists. Another is the supposition that the limiting process

will converge on the same limit every time a series of experiments is run. Another is that we can reliably estimate how many experiments are needed to establish this probability.

Given these problems we put aside the intuitive definition and build one around simpler axioms. Define a **probability space** as a pair (S, P) consisting of a sample space S made of events E, and a function P defined on the events E. An event is, to be precise, a specific subset of the sample space; we can build an intuitive feel for it by calling it a possible outcome of some process, using the ordinary English senses of the words. The function P satisfies these axioms [371] [426].

Axiom 1 *For each* $E \subseteq S$, $0 \leq P\{E\} \leq 1$.

That is, probabilities are defined to be between zero and one (inclusively). The greater the probability the greater the chance an event occurs.

Axiom 2 $P\{S\} = 1$.

This is a difficult axiom to dispute; the outcome of an experiment must be some element of the sample space S.

Axiom 3 *If* E_1, E_2, E_3, $\cdots$, E_n *is any collection of mutually exclusive events, then* $P\{\cup_{i=1}^{n} E_i\} = \sum_{i=1}^{n} P\{E_i\}$.

That is, if it is not possible for several of a set of events to simultaneously occur, then the probability that exactly one of them does occur is the sum of the probabilities of any of them occurring.

Axiom 4 *If* E *and* F *are events and* α *a number between* 0 *and* 1*, then there exists an event* S *for which* $P\{S\} = \alpha$ *and for which* $P\{S \cap E\} = P\{S\}P\{E\}$ *and* $P\{S \cap F\} = P\{S\}P\{F\}$.

That is, whatever events we have we can assume the existence of other independent events which are as probable or improbable as desired.

From these axioms[1] we are able to build the studies of probability familiar to every student, as well as the properties used for statistical mechanics. Trusting that the ideas of independent events and of conditional probability are familiar we move on to random variables.

2.3 Random Variables and Distribution Functions

The sample space S is if nothing else a set. We can define a function from S to the real numbers. Such a real-valued function is known as a **random variable**.

[1] There is one more assumption we must make, if the sample space has infinitely many elements – we must assume that P is defined only for events which are measurable in the sample space. This restriction will not impair our work, but it is needed for the accurate analysis of infinitely large sets.

For example, if the experiment is tossing ten fair coins, the sample space is all 2^{10} possible ways the coins may fall, and a random variable of interest might be the number of heads which turn up. If the experiment is rolling a pair of dice, the sample space is the set of 36 combinations of outcomes of the dice; our random variable is the sum of the numbers on those dice.

These variables are interesting because the typical view of statistical mechanics is that detailed information of a system is not important – what we want to know are properties like energy or entropy, which are random variables by this definition. We will move towards first the distribution functions of a random variable X – the probability of X falling within a specified range – and the expectation value – the mean value of X.

Define the **cumulative distribution function** F of the random variable X to be probability that X is less than or equal to d:

$$F(d) = P\{X \leq d\} \tag{2.2}$$

this function satisfies several properties; among them [58] [371] [426]:

1. F is non-decreasing: if $a < b$, then $F(a) \leq F(b)$.
2. $\lim_{d\to\infty} F(d) = 1$.
3. $\lim_{d\to-\infty} F(d) = 0$.
4. F is continuous from the right: for any decreasing sequence d_n which converges to d, $\lim_{n\to\infty} F(d_n) = F(d)$. (Why is F not necessarily continuous?)

Define the **probability mass function** $p(x)$ of X to be

$$p(x) = P\{X = x\} \tag{2.3}$$

and satisfying the condition X must take on one of the values x_i:

$$\sum_{i=1}^{\infty} p(x_i) = 1 \tag{2.4}$$

The cumulative distribution function $F(c)$ we construct by letting

$$F(c) = \sum_{\text{all } x\leq c} p(x) \tag{2.5}$$

which it is easy to verify satisfies the above properties.

We are not interested only in discrete random variables. Continuous random variables such as the energy[2] of a system are handled similarly to the discrete case.

[2] The energy of a real dynamical system may be drawn from an interval. In the numerical simulations we do, we confine ourselves to the floating point number system and so will use what is "really" a discrete system. Generally though we will use the terminology of continuous random variables.

Let X be a random variable which may have any value from an uncountably infinite set of values. This X is a **continuous random variable** if there exists a function f, defined on all real numbers and non-negative, for which

$$P\{X \in A\} = \int_A f(x)dx \tag{2.6}$$

for every subset A of the reals. We also require of f that

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{2.7}$$

as $\int_{-\infty}^{\infty} f(x)dx$ is the probability that X is some real number. This function f is the **probability density function** of X.

We may go from the probability density function to the cumulative distribution function by setting

$$F(c) = P\{X \in (-\infty, c]\} \tag{2.8}$$

$$= \int_{-\infty}^{c} f(x)dx \tag{2.9}$$

Equivalently,

$$\frac{d}{dc}F(c) = f(c) \tag{2.10}$$

and the probability of observing X to be within the range (c, d) is equal to

$$P\{X \in (-\infty, d]\} - P\{X \in (-\infty, c]\} = \int_{-\infty}^{d} f(x)dx - \int_{-\infty}^{c} f(x)dx \tag{2.11}$$

$$= \int_{c}^{d} f(x)dx \tag{2.12}$$

The probability density function is an approximation of the probability that, for any c, that X will be near c. Consider an interval of width ϵ centered around c:

$$P\{c - \frac{\epsilon}{2} \leq X \leq c + \frac{\epsilon}{2}\} = \int_{c-\frac{\epsilon}{2}}^{c+\frac{\epsilon}{2}} f(x)dx \tag{2.13}$$

$$\approx \epsilon f(c) \tag{2.14}$$

if f is continuous and ϵ sufficiently small.

2.4 Expectation Values and Averages

Given a random variable – the sum of a pair of dice, the distance a car may travel without needing refuelling, the kinetic energy of a gas – it is hard to

avoid wondering what its average value is. The word "average" has multiple meanings; ours is a weighted arithmetic mean. Values of X more likely to occur should weigh more in the average than improbable ones. The weighted mean we finally study is called the expectation value of X.

Define the **expectation value** E of discrete variable X to be [371] [426]

$$E[X] = \sum_i x_i P\{X = x_i\} \tag{2.15}$$

and for a continuous variable X to be

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx \tag{2.16}$$

(which are the same formulas used to find the center of mass for a set of discrete masses and for a continuous mass, a similarity reflected in the terms "probability mass" and "probability density" functions).

This expectation value is the "average" which matches our intuitive expectation: if we were to run a large number of experiments and measure X for each of them, the mean of these measured values of X will tend towards $E[X]$.

Given the real-valued function g we can define a new random variable $Y = g(X)$. If we know the probability mass or density function for X, do we know the expectation value for Y?

Proposition 1. *Suppose X is discrete; then*

$$E[Y] = E[g(X)] = \sum_i g(x_i) p(x_i) \tag{2.17}$$

Suppose X is continuous; then

$$E[Y] = E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx \tag{2.18}$$

Proof. Notice that there is for each y_j a set of points G_j, such that for any x_i in G_j we have $g(x_i) = y_j$. This lets us reorganize, in the discrete case, the sum over all the different values x_i into a sum over all the different values y_j of $g(x_i)$. So:

$$\sum_i g(x_i) p(x_i) = \sum_j \sum_{x_i \in G_j} g(x_i) p(x_i) \tag{2.19}$$

$$= \sum_j \sum_{x_i \in G_j} y_j p(x_i) \tag{2.20}$$

$$= \sum_j y_j \sum_{x_i \in G_j} p(x_i) \tag{2.21}$$

$$= \sum_j y_j P\{g(X) = y_j\} \tag{2.22}$$

$$= E[g(X)] \tag{2.23}$$

□

The proof of the continuous case is similar in inspiration, although to use the same outline requires the Lebesgue[3]-Stieltjes[4] definition of an integral. To avoid introducing that much supporting material we will instead narrow the focus so that g is a nonnegative function, and introduce a lemma.

Lemma 1. *For a nonnegative continuous random variable Y,*

$$E[Y] = \int_0^\infty P\{Y \geq y\}dy \tag{2.24}$$

Proof. Let f_Y be the probability density function for Y. Therefore $P\{Y > y\} = \int_y^\infty f_Y(x)dx$. And (notice the swapping the order of integration on the second line of this derivation)

$$\int_0^\infty P\{Y > y\}dy = \int_0^\infty \int_y^\infty f_Y(x)dxdy \tag{2.25}$$

$$= \int_0^\infty \left(\int_0^x dy\right) f_Y(x)dx \tag{2.26}$$

$$= \int_0^\infty x f_Y(x)dx \tag{2.27}$$

$$= E[Y] \tag{2.28}$$

Now we can complete the proof of the restricted form of Proposition 1. Let g be any continuous nonnegative function. For any y, let G_y be the set of x for which $g(x) > y$. We have

$$E[g(X)] = \int_0^\infty P\{g(X) > y\}dy \tag{2.29}$$

$$= \int_0^\infty \int_{G_y} f(x)dxdy \tag{2.30}$$

$$= \int_{G_0} \left(\int_0^{g(x)} dy\right) f(x)dx \tag{2.31}$$

$$= \int_{G_0} g(x)f(x)dx \tag{2.32}$$

[3] Henri Léon Lebesgue, 1875 - 1941, was one of the founders of measure theory, although he did not concentrate on that field. He found it too general a theory for his tastes and preferred to work on smaller, specific topics. [288]

[4] Thomas Jan Stieltjes, 1856 - 1894, besides extending the definition of the integral, is regarded as the founder of analysis of continued fractions, numbers defined by $1/(a + 1/(b + 1/(c + 1/(\cdots))))$ for some integer sequence a, b, c, d, $\cdots$. His time as a student was spent reading Gauss and Jacobi, rather than attending lectures, which gave him a good background but made him fail his exams. [288]

□

The expectation value $E[g(X)]$ is in general *not* equal to $g(E[X])$. In fact the expectation value is – as one might guess from observing that it is either a sum or an integral – linear; $E[g(X)]$ will only equal $g(E[X])$ if the function g is of the form $a \times X + b$.

Proposition 2. *$E[aX + b] = aE[X] + b$ for constants a, b.*

Proof. When X is a discrete random variable,

$$E[aX + b] = \sum_i (ax_i + b)p(x_i) \tag{2.33}$$

$$= a\sum_i x_i p(x_i) + b\sum_i p(x_i) \tag{2.34}$$

$$= aE[X] + b \tag{2.35}$$

When X is a continuous random variable,

$$E[aX + b] = \int_{-\infty}^{\infty} (ax + b)f(x)dx \tag{2.36}$$

$$= a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx \tag{2.37}$$

$$= aE[X] + b \tag{2.38}$$

□

2.5 Variance

Though the expectation value of a variable describes some of its behavior, it is not the only relevant quantity. A variable X which is always equal to zero has the same expectation value as a variable Y which is any integer between -5 and 5 with uniform probability. As a measure of how distributed the numbers are we can introduce (analogously to the moment of inertia of a set of mass) the **moment** of the variable X. For any n, the n**th moment** of X is

$$E[X^n] = \sum_i x_i^n p(x_i) \tag{2.39}$$

if X is discrete and

$$E[X^n] = \int_{-\infty}^{\infty} x_i^n f(x)dx \tag{2.40}$$

if X is continuous (and assuming the expectation value exists).

These moments are dependent on the values of X; if we define a new variable $Y = X + c$ for a constant c the moments of Y will be different from

those of X though most would say the distribution of X is the same as that of Y. We can restore this "translation invariance" by using μ_X, the expectation value of X, as a reference point. Define the n**th central moment** of X to be

$$E[(X - \mu_X)^n] = \sum_I (x_i - \mu_X)^n p(x_i) \tag{2.41}$$

for the discrete case and

$$E[(X - \mu_X)^n] = \int_{-\infty}^{\infty} (x - \mu_X)^n f(x) dx \tag{2.42}$$

for the continuous (and again providing the integrals exist) [371].

The second central moment of X is known as the **variance**, $Var[X]$, and is quite often used. Its square root is known as the **standard deviation**, $\sigma[X]$. Typically the easiest way to calculate the variance is to use the second moment of X and the square of the first moment of X [371] [426].

Proposition 3. $Var[X] = E[X^2] - \mu_X^2$.

Proof. Since the variance of X is its second central moment we have

$$\begin{aligned} Var[X] &= E[(X - \mu_X)^2] && (2.43)\\ &= E[X^2 - 2\mu_X X + \mu_X^2] && (2.44)\\ &= E[X^2] - 2\mu_X E[X] + \mu_X^2 && (2.45)\\ &= E[X^2] - 2\mu_X^2 + \mu_X^2 && (2.46)\\ &= E[X^2] - \mu_X^2 && (2.47) \end{aligned}$$

As $\mu_X = E[X]$ then $Var[X] = E[X^2] - (E[X])^2$. The variance or the standard deviation most often describe whether the values of X are distributed close together (a small variance) or widely apart (a large variance).

□

Proposition 4. $Var[aX + b] = a^2 Var[X]$

Proof. Remembering proposition 2 and the definition of variance,

$$\begin{aligned} Var[aX + b] &= E[(aX + b - a\mu_X - b)^2] && (2.48)\\ &= E[(aX - a\mu_X)^2] && (2.49)\\ &= E[a^2(X - \mu_X)^2] && (2.50)\\ &= a^2 E[(X - \mu_X)^2] && (2.51)\\ &= a^2 Var[X] && (2.52) \end{aligned}$$

□

2.6 Multiple Variables and Independence

To this point we have considered only a single random variable X. It is almost inevitable we will want to measure several quantities in experiments, so we want to establish our probability tools for multiple variables.

Suppose that we have random variables X and Y. The **cumulative joint probability distribution function** is defined to be

$$\begin{aligned} F(c,d) &= P\{X \leq c, Y \leq d\} && (2.53)\\ &= P\{X \in (-\infty, c] \cap Y \in (-\infty, d]\} && (2.54) \end{aligned}$$

We also define the **marginal distributions** of X and Y, which examine the probability for a single variable assuming the other is completely free. The marginal distribution of X is $F_X(c)$ and equals

$$\begin{aligned} F_X(c) &= P\{X \leq c\} && (2.55)\\ &= P\{X \leq c, Y < \infty\} && (2.56)\\ &= \lim_{d\to\infty} P\{X \leq c, Y \leq d\} && (2.57)\\ &= \lim_{d\to\infty} F(c,d) && (2.58)\\ &= F(c,\infty) && (2.59) \end{aligned}$$

and similarly

$$\begin{aligned} F_Y(d) &= P\{Y \leq d\} && (2.60)\\ &= P\{X < \infty, Y \leq d\} && (2.61)\\ &= \lim_{c\to\infty} P\{X \leq c, Y \leq d\} && (2.62)\\ &= \lim_{c\to\infty} F(c,d) && (2.63)\\ &= F(\infty,d) && (2.64) \end{aligned}$$

There is, for the case of discrete random variables, a **joint probability mass function** defined by

$$p(c,d) = P\{X = c, Y = d\} \tag{2.65}$$

with separate **probability mass functions** p_X and p_Y defined by

$$\begin{aligned} p_X(c) &= P\{X = c\} && (2.66)\\ &= \sum_d p(c,d) && (2.67)\\ p_Y(d) &= P\{Y = d\} && (2.68)\\ &= \sum_c p(c,d) && (2.69) \end{aligned}$$

Similarly for continuous random variables the **joint probability density function** $f(x, y)$ is

$$P\{X \in C, Y \in D\} = \int_D \int_C f(x, y) dx dy \tag{2.70}$$

with separate **probability density functions** f_X and f_Y defined by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \tag{2.71}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx \tag{2.72}$$

We have mentioned independent random variables. We need now to give this idea a precise definition, as the introduction of several variables makes it less obvious independence could be taken for granted. The random variables X and Y are said to be **independent** if for any two subsets C and D of the probability space,

$$P\{X \in C, Y \in D\} = P\{X \in C\} P\{Y \in D\} \tag{2.73}$$

This is equivalent to showing that for all c and d,

$$P\{x \leq c, y \leq d\} = P\{x \leq c\} P\{y \leq d\} \tag{2.74}$$

Two variables are **dependent** if they are not independent. That is, they are independent if the events $E_C = \{X \in C\}$ and $E_D = \{Y \in D\}$ are independent.

Proposition 5. *If X, Y have joint probability density function $f(x, y)$ then they are independent if and only if*

$$f(c, d) = f_X(c) f_Y(d) \tag{2.75}$$

Another tool useful to examine independence is the **covariance**, a generalization of the variance of a single variable. For X and Y define the covariance $Cov[X, Y]$ to equal $E[(X - \mu_X)(Y - \mu_Y)]$, where $\mu_X = E[X]$ and $\mu_Y = E[Y]$.

Proposition 6. *If X and Y are independent then $Cov[X, Y] = 0$.*

Proof. We start with a lemma relating the covariance of X and Y to the expectation values of X, Y, and their product XY.

Lemma 2.

$$Cov[X, Y] = E[XY] - E[X]E[Y] \tag{2.76}$$

Proof.

$$Cov[X,Y] = E[XY - \mu_X Y - X\mu_Y + \mu_X\mu_Y] \tag{2.77}$$
$$= E[XY] - \mu_X E[Y] - \mu_Y E[X] + \mu_X\mu_Y \tag{2.78}$$
$$= E[XY] - \mu_X\mu_Y. \tag{2.79}$$

□

Returning to the proof of Proposition 5, we examine the expectation value of the product XY.

$$E[XY] = \sum_j \sum_i x_i y_i P\{X = x_i, Y = y_i\} \tag{2.80}$$
$$= \sum_j \sum_i x_i y_i P\{X = x_i\} P\{Y = y_i\} \tag{2.81}$$
$$= \left(\sum_j y_j P\{Y = y_j\}\right)\left(\sum_i x_i P\{X = x_i\}\right) \tag{2.82}$$
$$= E[Y]E[X] \tag{2.83}$$

So by Lemma 2,

$$Cov[X,Y] = 0. \tag{2.84}$$

□

2.7 Limiting Theorems

Next we look for assurance a finite number of experiments can measure something meaningful: will sufficiently many experiments approximate the "infinite" continuous limit? The **central limit theorem** shows the result of a sufficiently large number of independent, identically distributed random variables (sometimes abbreviated as "iid random variables") will approach the familiar Gaussian[5] bell curve or "normal" distribution of a random variable [58] [371] [426].

We begin with Markov's [6] Inequality. This is the first of two propositions which let us use only the mean and variance of a probability distribution to establish bounds on the probability of events.

[5] Johann Carl Friedrich Gauss, 1777 - 1855, is another person hard to exaggerate. He developed important principles for every field from probability to non-Euclidean geometry to mechanics; he lead a geodesic survey of the then-independent state of Hannover, and developed the least-squares approximation method in the hope of finding the asteroid Ceres, discovered 1 January 1801 by Giuseppe Piazzi, 1746-1826, and lost shortly after. His predictions were right. [288]

[6] Andrei Andreevich Markov, 1856 - 1922, besides Markov sequences and probability, studied continued fractions, which were pioneered by his teacher Pafnuty Lvovich Chebyshev. [92]